

## **Projektbericht**

# Entwicklung eines Interfaces zur kognitiven Anreicherung der Interaktion von Lokführern mit der Punktförmigen Zugbeeinflussung

**Technische Universität Berlin**

Fachgebiet Mensch-Maschine-Systeme

Interdisziplinäre Projektarbeit WS 2013/14

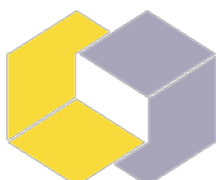
in Kooperation mit dem Deutschen Zentrum für Luft- und Raumfahrt e. V.

**Verfasser:** Peter Brandstädter, Osman Gürbüz, Theresa Jarchow,  
Irina Kuznetsova, Sophie Neef, Milenko Petrovic,  
Kevin Seiler, Christian Stropp

**Auftraggeberin:** Sonja Gieseemann, Deutsches Zentrum für Luft- und Raumfahrt e. V.

**Betreuerinnen:** M. Sc. Antje Venjakob, Technische Universität Berlin  
B. Sc. Assiel Afram, Technische Universität Berlin

**Vorgelegt am:** 14.02.2014



## Abstract

Die spezifischen Eigenschaften von Schienenfahrzeugen, wie lange Bremswege und keine Ausweichmöglichkeiten durch Schienenbindung, erfordern besondere Sicherheitsmaßnahmen. In Deutschland überwacht die Punktförmige Zugbeeinflussung (PZB) auf dem Großteil des Schienennetzes die Reaktion der Triebfahrzeugführer auf die angezeigten Signale. Dabei muss der Triebfahrzeugführer Vorsignale durch Betätigung einer Wachsamkeitstaste quittieren. Bleibt dies aus, erfolgt eine Zwangsbremung. Die Quittierungshandlung wird allerdings sehr automatisiert durchgeführt, so dass eine aufmerksame Verarbeitung der Signale nicht gewährleistet ist. Im Zuge der Projektarbeit sollte deshalb ein neues Interface entwickelt werden, welches eine Differenzierung der unterschiedlichen Vorsignalbegriffe sicherstellt. Es entstanden zwei Konzepte, ein Hebel- und ein Touch-Interface, für die Prototypen angefertigt und innerhalb eines experimentellen Settings durch vier Triebfahrzeugführer evaluiert wurden. Dabei wurde das Hebel-Interface als benutzerfreundlicher und weniger beanspruchend bewertet als das Touch-Interface. Diese Exploration liefert erste Hinweise für die Generierung weiterer Konzepte. Die beschränkte Möglichkeit zur Generalisierung durch eine sehr kleine Stichprobe und eine vereinfachte Experimentalumgebung wird diskutiert.

The specific properties of railed vehicles, like a long breaking distance and the fixation to rails, require special safety measures. In Germany the Punktförmige Zugbeeinflussung (intermittent automatic train running control) monitors the reaction of the train operator on given railway signals. The train operator has to acknowledge certain signals with a push button or the train will be forced to stop. Because the acknowledgement operation is highly automated and can be extended to all signals, the attentive processing of the railway signals can not be ensured. This project work served to develop new concepts for a new interface, which allows for differentiation between different states of distant signals. Two interface concepts – a lever and a touch based interface – were created, implemented as prototypes and evaluated by four train operators within an experimental setting. The lever was rated to be more usable and less demanding in terms of workload than the touch interface. Cues from this exploration can be used to derive further concepts. Limits in generalization of the results due to a small sample size and simplified experimental environment are discussed.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>4</b>
<b>2</b>	<b>Problemstellung</b>	<b>5</b>
2.1	Aufgabenstellung	5
<b>3</b>	<b>Theoretischer Hintergrund</b>	<b>6</b>
3.1	Punktförmige Zugbeeinflussung	6
3.2	Zugsicherungen in anderen Ländern	7
3.3	Anforderungen an den Triebfahrzeugführer	9
3.4	Interfacegestaltung	11
3.5	Methoden der Evaluation	12
<b>4</b>	<b>Konzeption und prototypische Umsetzung</b>	<b>13</b>
4.1	Hebel	13
4.2	Touchscreen	15
<b>5</b>	<b>Methode</b>	<b>17</b>
5.1	Stichprobe	17
5.1.1	Rekrutierung	17
5.1.2	Beschreibung der Stichprobe	17
5.2	Versuchsdesign	17
5.2.1	Materialien	18
5.2.2	Versuchsaufbau	19
5.2.3	Versuchsablauf	20
5.3	Auswertungsmethoden	21
<b>6</b>	<b>Ergebnisse</b>	<b>22</b>
6.1	Lautes Denken	22
6.2	Interview	24
6.3	Fehlerrate	26
6.4	Subjektive Beanspruchung (NASA-TLX)	27
6.5	Subjektive Usability (ISONORM 9241/110-S)	30
<b>7</b>	<b>Diskussion</b>	<b>33</b>
<b>8</b>	<b>Ausblick</b>	<b>35</b>
<b>9</b>	<b>Appendix</b>	<b>38</b>
A.	Anhang – Demografischer Fragebogen	38
B.	Anhang – Einführungs und Begrüßungstext	40
C.	Anhang – Lautes Denken: Ablauf	42
D.	Anhang – Lautes Denken: Text zum Vorlesen	44
E.	Anhang – Instruktionen für die Versuchsperson	45
F.	Anhang – NASA-TLX	47
G.	Anhang – ISONORM 9241/110-S	49
H.	Anhang – Interview	53
	<b>Tabellenverzeichnis</b>	<b>56</b>
	<b>Abbildungsverzeichnis</b>	<b>57</b>
	<b>Literaturverzeichnis</b>	<b>58</b>

# 1 Einleitung

Heutzutage ist die Bahn ein allgemein und häufig genutztes Verkehrsmittel, so beförderte die Deutsche Bahn im Jahresverlauf 2012 ca. 1.92 Milliarden Menschen (FOCUS Online). Ein Grund dafür könnte sein, dass Züge als besonders sicher für den Weg zur Schule, zur Arbeit oder in den Urlaub gelten. Doch woran liegt es, dass der Zugverkehr so sicher ist? Ein Grund dafür ist, dass die Lokführer mehrfach in ihren Handlungen kontrolliert werden. Dies geschieht in Deutschland z. B. durch die Punktförmige Zugbeeinflussung, die an kritischen Punkten auf der Strecke den Zug zwangsbremsen kann. Sie soll verhindern, dass zwei Züge auf denselben Streckenabschnitt fahren und damit vor Kollisionen schützen.

In den letzten Jahren kam es in den europäischen Ländern zu mehreren schweren Unfällen, bei denen Signale durch den Lokführer missachtet wurden. Die punktförmigen Zugbeeinflussungssysteme, welche länderspezifisch technische Unterschiede aufweisen, aber sich in der Funktionsweise und damit der dargestellten Problematik stark ähneln, waren auf den betreffenden Streckenabschnitten oft nicht installiert oder reichten schlicht nicht aus, um die kritische Situation zu verhindern.

Punktförmige Zugbeeinflussungssysteme sind aus technischer Perspektive eine logische und hilfreiche Maßnahme, um den Bahnbetrieb sicherer zu machen. Allerdings treten Probleme auf, wenn der Mensch bei der Interaktion mit dem technischen System ein falsches Verständnis dafür entwickelt und die Bedienung des Systems nicht an den äußeren Kontext angepasst ist. Es besteht die Notwendigkeit, die Schnittstelle zwischen Mensch und Maschine einer besonderen psychologischen Untersuchung zu unterziehen und neue Konzepte zu entwickeln, die insgesamt zu einer höheren Sicherheit beitragen können.

## 2 Problemstellung

Die Punktförmige Zugbeeinflussung (PZB) arbeitet nach einfachen Regeln, die nicht alle kritischen Situationen ausschließen können. Auf vereinzelt Streckenabschnitten ist sie nicht vorhanden. Die Lokführer dürfen sich also nicht auf dieses System verlassen und müssen stets wachsam bleiben. Das Deutsche Zentrum für Luft- und Raumfahrt (DLR) stellte im Rahmen der Forschung fest, dass die starke Gewöhnung der Lokführer an die PZB insgesamt zu einer Fehleinschätzung des Systems führt und daher von ihnen oft fälschlicherweise als Assistenzsystem beschrieben wird.

Es bedarf weiterhin der korrekten Handlungsausführung durch die Lokführer, die hiermit als Menschen ein potenzielles Risiko im System darstellen. Im Regelbetrieb kann es zu Monotonie und Ermüdung aufgrund der erforderlichen Daueraufmerksamkeit sowie zu Ablenkung kommen. In solchen Situationen begünstigen das übersteigerte Vertrauen in die PZB, die zurzeit sehr einfach gestaltete Bedienung des Systems und das Fehlen der Information, ob die PZB auf einem Streckenabschnitt aktiv ist, Fehler, die fatale Folgen haben können.

### 2.1 Aufgabenstellung

Im Rahmen ihrer Forschungsarbeit am Deutschen Zentrums für Luft- und Raumfahrt verfasste Sonja Giesemann die Aufgabenstellung, ein neues Interface zur kognitiven Anreicherung der Interaktion von Lokführern mit der Zugsicherung zu entwickeln. Die antrainierte Reaktion der Lokführer auf ein Vorsignal in Form der Quittierung durch Tastendruck soll durch eine Handlung ersetzt werden, die die aufmerksame Verarbeitung des Vorsignals überprüft. Dabei soll das System der Punktförmigen Zugbeeinflussung als technische Basis bestehen bleiben.

Für die verschiedenen Vorsignalbegriffe sollten Metaphern entwickelt werden, die sich eignen, bei der Quittierung eine tiefere Verarbeitung zu erzwingen. Auch den unterschiedlichen Modalitäten, in denen die Quittierungshandlung vollzogen und Feedback darüber präsentiert werden kann, sollte Beachtung geschenkt werden.

Ziel des Projekts war es, in einem nutzerzentrierten Prozess zwei Interface-Konzepte zu entwickeln, sie prototypisch umzusetzen und mit ausgebildeten Triebfahrzeugführern zu evaluieren.

### 3 Theoretischer Hintergrund

#### 3.1 Punktförmige Zugbeeinflussung

Unter dem Punktförmigen Zugbeeinflussungssystem (PZB) wird bei den Eisenbahngesellschaften in Deutschland eine ihrer wichtigsten Sicherungseinrichtungen verstanden. Falls ein Triebwagenführer mit seinem Zug in eine gesperrte Strecke einfährt oder auch in einem Streckenabschnitt die vorgeschriebene Geschwindigkeiten (werden über Sensoren gemessen) an Gefahrenpunkten nicht einhält, kann durch die PZB die Einleitung einer Zwangsbremmung erfolgen, um so Unfälle von Zügen und Gefährdungen von Menschenleben zu vermeiden. Vor immer mehr Vor- bzw. Hauptsignalen, Bahnübergangssignalen, Sperrsignalen und Einfahrten in Stumpfgleise haben die Bahngesellschaften inzwischen in diese Sicherheitstechniken Milliarden von Euro investiert. Dabei handelt es sich im Prinzip um eine relativ einfache aber sehr wirksame Technik. (Was ist ein Punktförmiges Zugbeeinflussungssystem kurz PZB, 2013)

Auf der Höhe des Schienenkopfes, rechts in Fahrtrichtung, werden Indusimagnete, jeweils nach Standort mit passiven Schwingkreisen (Frequenz 500, 1000, 2000 Hz) angebaut. Das entsprechende Gegenstück, mit einem elektrischen Schwingkreis, ist an dem Zug installiert.

Das Triebfahrzeug wird durch den 1000 Hz-Magnet beeinflusst, wenn der Triebfahrzeugführer während seiner Fahrt auf einen Vorsignal mit „Halt erwarten“ oder „Langsamfahrt erwarten“ stößt. Anschließend muss die Wachsamkeitstaste innerhalb von 4 Sekunden betätigt werden, somit bestätigt er dem System, dass er das Signal wahrgenommen hat. Wird dies nicht betätigt, ertönt ein Warnsignal und es erfolgt eine Zwangsbremmung.

(Grundlagen der Stellwerkstechnik, 2004)

Wird bei der Fahrt ein 500-Hz-Magnet, zwischen dem Vor- und Hauptsignal, überfahren bzw. ausgelöst, kommt es zu einer Kontrolle der Geschwindigkeit. Der 2000-Hz-Magnet befindet sich am Hauptsignal und sorgt dafür, dass an einem Halt zeigenden Signal, eine sofortige Zwangsbremmung ausgelöst wird. Es befindet sich auf dem Befehlsstand die Befehlstaste, mit dem der Triebfahrzeugführer damit eine 2000-Hz-Beeinflussung überbrücken kann, wenn z.B. ein defektes Hauptsignal mit Ersatzsignal oder von dem Fahrdienstleiter ein schriftlicher Befehl vorliegt. (PZB 90, 2011) In der *Abbildung 1* ist das Betriebs-

programm grafisch dargestellt (PZB 90 Betriebsprogramm.PNG - Wikimedia Commons, 2013).

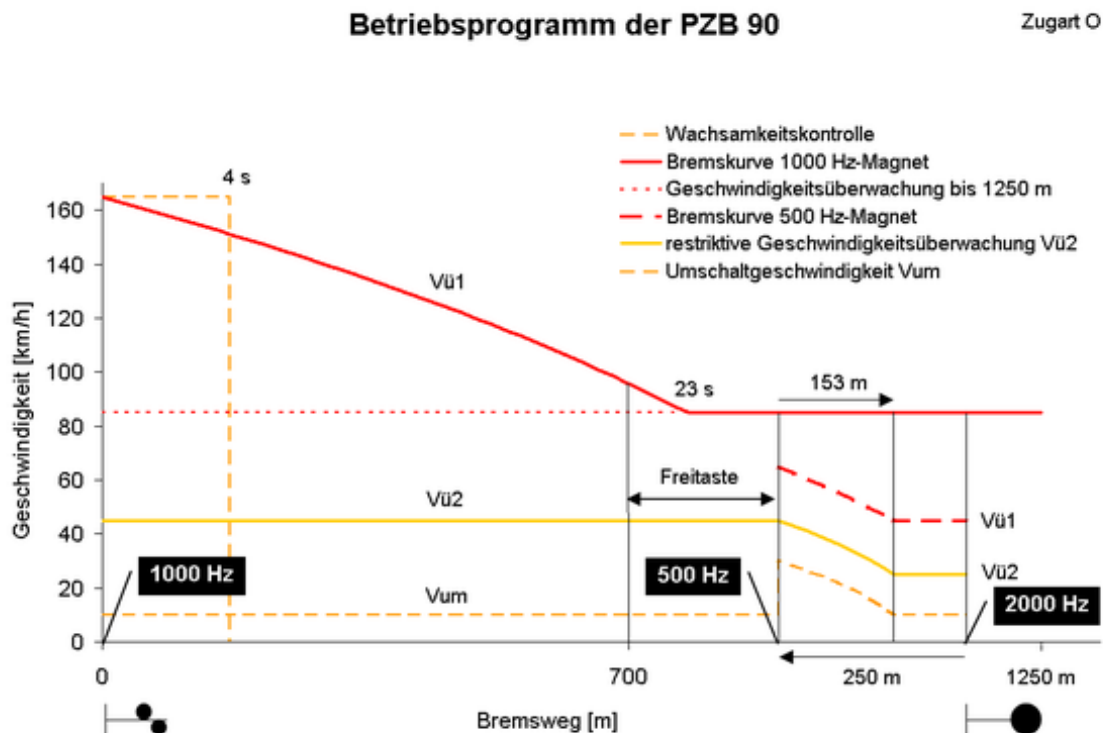


Abbildung 1: Betriebsprogramm der PZB

### 3.2 Zugsicherungen in anderen Ländern

Weltweit gibt es eine Vielzahl von Zugbeeinflussungssystemen, die jeweils unterschiedlich funktionieren. Im Folgenden werden die in Europa am weitesten verbreiteten Zugbeeinflussungssysteme aufgeführt.

Wie *Tabelle 1* zu entnehmen ist, sind in Europa unterschiedliche Zugsicherungssysteme im Einsatz. Durch die Vielzahl verschiedener Systeme wird die Interoperabilität im europäischen Raum beeinträchtigt. Um dieser Beeinträchtigung im grenzüberschreitenden Schienenverkehr entgegen zu wirken, wird derzeit von zahlreichen europäischen Bahnbetreibern an der Entwicklung eines gemeinsamen Zugsicherungssystems gearbeitet, genannt ETCS (*European Train Control Center*). Bei diesem Zugsicherungssystem werden dem Lokführer die Fahrerlaubnis, Geschwindigkeitsangaben und Streckendaten im Führerstand angezeigt. Deshalb kann auf eine Außensignalisierung verzichtet werden. Die Überwachung der

unterschiedlichen Zugparameter wird durch die Streckenzentrale und die des ETCS-Fahrzeugrechners gewährleistet. Hierbei gibt es bei der Signalübertragung unterschiedliche Arten. Die Vorgaben der EU wurden von den Mitgliedsstaaten in nationales Recht überführt und die geforderte Umsetzung ist bereits eingeleitet. (DB Netz AG , 2012)

**Tabelle 1: Systeme mit punktförmiger Zugbeeinflussung in anderen Ländern**

Systembezeichnung	Übertragungssystem	Land
Indusi, PZB 90	induktives Resonanzsystem	Deutschland, Österreich
Crocodile	Schleifkontakte	Frankreich, Belgien, Luxemburg
Signum	magnetisches System	Schweiz
ZUB 121	Transponder, Kurzscheifen	Schweiz
ZUB 123	Transponder, Kurzscheifen	Dänemark
TBL1, TBL2	Transponder	Belgien
KVB	Transponder	Frankreich
AWS	magnetisches System	Großbritannien
ASFA	induktives Resonanzsystem	Spanien
EBICAB2	Transponder	Norwegen, Schweden
L 10 000	Transponder	Schweden



### 3.3 Anforderungen an den Triebfahrzeugführer

Die verkehrsmedizinische und psychologische Eignungsbeurteilung ist gesetzlich verankert und verfolgt das Ziel, Menschen entsprechend den sicherheitsrelevanten Anforderungen der Tätigkeit Triebfahrzeugführer einzusetzen. Als rechtliche Grundlage dient das Allgemeine Eisenbahngesetz (Allgemeines Eisenbahngesetz (AEG) - Gesetze im Internet, 2005) sowie die „Verordnung über die Erteilung der Fahrberechtigung an Triebfahrzeugführer sowie die Anerkennung von Personen und Stellen für die Ausbildung und Prüfung“ (Triebfahrzeugführerscheinverordnung - TfV - Gesetze im Internet, 2011).

Anlage 4 der TfV (Anlage 4 TfV - Buzer.de, 2011) enthält die genauen Anforderungen der Einstellungsuntersuchungen für den Triebfahrzeugführer.

Ein Triebfahrzeugführer darf nicht unter gesundheitlichen Störungen leiden, und Arzneimittel oder Stoffe nehmen, die insbesondere Folgendes auslösen können: plötzliche Bewusstlosigkeit; Verminderung der Aufmerksamkeit oder der Konzentration; plötzliche Handlungsunfähigkeit; Verlust des Gleichgewichts oder der Koordination; erhebliche Einschränkung der Mobilität. (1.1)

Folgende Anforderungen an das Sehvermögen müssen erfüllt sein: Fern-Sehschärfe mit oder ohne Sehhilfe: 1,0; mindestens 0,5 für das schlechtere Auge; maximale Korrekturlinsenstärke: Hyperopie +5 / Myopie -8; Abweichungen sind in Ausnahmefällen zulässig; eine Entscheidung erfolgt durch die zuständige Behörde im Rahmen der Erteilung des Triebfahrzeugführerscheins nach Einholung einer Stellungnahme eines Augenarztes; Sehvermögen nahe und mittlere Entfernung: ausreichend, mit oder ohne Sehhilfe; Kontaktlinsen und Brillen sind zulässig, sofern das Sehvermögen regelmäßig von einem Augenarzt überprüft wird; normale Farbwahrnehmung: Verwendung eines anerkannten Tests wie des Ishihara-Tests; Sichtfeld: vollständig; Sehvermögen beider Augen: effektiv; nicht erforderlich, wenn der Betreffende über eine angemessene Anpassung und ausreichende Kompensationserfahrung verfügt; nur erforderlich, wenn der Betreffende das binokulare Sehvermögen nach Aufnahme der Tätigkeit verloren hat; binokulares Sehvermögen: effektiv; Erkennen farbiger Signale: die Prüfung erfolgt auf der Grundlage der Erkennung einzelner Farben, nicht auf der Grundlage relativer Unterschiede; Kontrastempfindlichkeit: gut; keine fortschreitenden Augenkrankheiten; Linsenimplantate, Keratotomien und Keratektomien sind nur zulässig, wenn sie jährlich oder in vom Arzt festgelegten regelmäßigen Abständen überprüft werden; keine Überempfindlichkeit gegen Blendung; farbige Kontaktlin-

sen und fotochromatische Linsen sind nicht zulässig, Linsen mit UV-Filter sind zulässig.  
(1.2)

Anforderungen an das Hör- und Sprachvermögen: Ausreichendes, durch ein Audiogramm nachgewiesenes Hörvermögen für ein Telefongespräch und die Fähigkeit, akustische Warnsignale und Funkmeldungen zu hören. Dafür gelten folgende Richtwerte: Es darf kein Hördefizit von über 40 dB bei 500 und 1.000 Hz vorliegen; es darf kein Hördefizit von über 45 dB bei 2.000 Hz bei dem Ohr, das die schlechtere Schallleitung aufweist, vorliegen; keine Anomalie des Vestibularapparats; keine chronische Sprachstörung auf Grund der Notwendigkeit, Mitteilungen laut und deutlich auszutauschen; die Verwendung von Hörhilfen ist in bestimmten Fällen zulässig. (1.3) Weiterhin regelt die TfV einen Mindestinhalt der Einstellungsuntersuchung: Zum einen ärztliche Untersuchungen (allgemeine ärztliche Untersuchung; Untersuchung der sensorischen Funktionen: Sehvermögen, Hörvermögen, Farbwahrnehmung; Blut- oder Urinalysen, um unter anderem eine eventuelle Zuckerkrankheit festzustellen, soweit sie zur Beurteilung der körperlichen Eignung des Bewerbers erforderlich sind; Ruhe-Elektrokardiogramm (EKG); Untersuchung auf psychotrope Stoffe wie beispielsweise verbotene Drogen oder psychotrope Arzneimittel sowie auf Alkoholmissbrauch, die die berufliche Eignung in Frage stellen.) (2.1) und zum anderen psychologische Untersuchungen (kognitive Fähigkeiten: Aufmerksamkeit und Konzentration, Gedächtnis, Wahrnehmungsfähigkeit, Urteilsvermögen; Kommunikation; psychomotorische Fähigkeiten: Reaktionsgeschwindigkeit, Koordination der Hände; tätigkeitsrelevante Persönlichkeits- und Einstellungsfaktoren.) (2.2)

Trotz der im Vorfeld durchgeführten Eignungstests und Untersuchungen hinsichtlich Aufmerksamkeit und Konzentration kommt es im Bahnbetrieb häufig zu Aufmerksamkeitsdefiziten aufgrund von Monotonie, Unterforderung und Ermüdung. Ein Grund für Aufmerksamkeitsdefizite während des Betriebes ist der geringe Handlungsspielraum bei der Fahraufgabe (Überwachung der Systeme, Fahrweg beobachten, Fahrpläne anwenden, Signale quittieren, bremsen oder beschleunigen).

### 3.4 Interfacegestaltung

Bei einem Interface handelt es sich um eine Benutzeroberfläche zwischen Mensch und Maschine. Die Gestaltung eines Interfaces, muss gewährleisten, dass der Benutzer seine Aufgaben effektiv, effizient und zufriedenstellend ausüben kann. Für die Konzeption liegt eine Vielzahl von Richtlinien vor. Dabei werden formale Bewertungskriterien, wie beispielsweise Ausrichtung, Anzahl oder Konsistenz von funktionalen Bewertungskriterien, wie Konsistenz, Farbgebung oder Erreichbarkeit unterschieden. Letztere basieren auf Erkenntnissen der ergonomischen Grundlagenforschung und sollen eine beeinträchtigungsfreie Ausübung von Aufgaben gewährleisten. Für den Entwicklungsprozess sollte stets beachtet werden, dass die Kriterien voneinander abhängen (Schmid & Maier, 2009).

Um die Eigenschaften und Grenzen des Benutzers sinnvoll in den Entwicklungsprozess einzubeziehen, sollte dieser stets im Fokus stehen. Mit der Interface-Gestaltung soll so eine problemlose und reibungslose Bedienung gewährleistet sein. Interfaces sollten deshalb so gestaltet sein, dass der Benutzer seine Aufgaben intuitiv und fehlerfrei ausführen kann.

Die nutzerfreundliche Gestaltung ist Teil eines internationalen Standards. Die Richtlinien zur Gestaltung der Mensch-Maschine-Interaktion sind in der DIN EN ISO 9241-110 (2008) beschrieben. In dieser Norm sind sieben Prinzipien der Dialoggestaltung verankert, in denen die Gestaltung von Geräten, die Einrichtung von Arbeitsplätzen oder die Definition von Arbeitsaufgaben beschrieben werden.

Es werden folgende Grundsätze beschrieben:

*Aufgabenangemessenheit:* “Ein interaktives System ist aufgabenangemessen, wenn es den Benutzer unterstützt, seine Arbeitsaufgabe zu erledigen, d. h., wenn Funktionalität und Dialog auf den charakteristischen Eigenschaften der Arbeitsaufgabe basieren, anstatt auf der zur Aufgabenerledigung eingesetzten Technologie. “

*Selbstbeschreibungsfähigkeit:* “Ein Dialog ist in dem Maße selbstbeschreibungsfähig, in dem für den Benutzer zu jeder Zeit offensichtlich ist, in welchem Dialog, an welcher Stelle im Dialog er sich befindet, welche Handlungen unternommen werden können und wie diese ausgeführt werden können.”

*Steuerbarkeit:* “Ein Dialog ist steuerbar, wenn der Benutzer in der Lage ist, den Dialogablauf zu starten sowie seine Richtung und Geschwindigkeit zu beeinflussen, bis das Ziel erreicht ist.”

*Erwartungskonformität:* "Ein Dialog ist erwartungskonform, wenn er den aus dem Nutzungskontext heraus vorhersehbaren Benutzerbelangen sowie allgemein anerkannten Konventionen entspricht."

*Fehlertoleranz:* "Ein Dialog ist fehlertolerant, wenn das beabsichtigte Arbeitsergebnis trotz erkennbar fehlerhafter Eingaben entweder mit keinem oder mit minimalem Korrekturaufwand seitens des Benutzers erreicht werden kann."

*Individualisierbarkeit:* "Ein Dialog ist individualisierbar, wenn Benutzer die Mensch-System-Interaktion und die Darstellung von Informationen ändern können, um diese an ihre individuellen Fähigkeiten und Bedürfnisse anzupassen."

*Lernförderlichkeit:* "Ein Dialog ist lernförderlich, wenn er den Benutzer beim Erlernen der Nutzung des interaktiven Systems unterstützt und anleitet."

Für die Konzeption benutzerfreundlicher Interfaces gilt es als empfehlenswert, sich an solchen Richtlinien zu orientieren.

### 3.5 Methoden der Evaluation

Im Rahmen nutzerzentrierter Entwicklung werden häufig zwei Gruppen von Evaluationsmethoden unterschieden: expertenbasierte und nutzerbasierte Methoden.

Grundlage expertenbasierter Evaluationsverfahren bilden Expertenurteile. Zu den expertenbasierten Evaluationsverfahren zählen zum Beispiel die heuristische Evaluation, die leitfadenbasierte Evaluation und die Methode des Cognitive Walkthrough. Bei der Methode der heuristischen Evaluation wird die Nutzerfreundlichkeit (Usability) einer Mensch-Maschine-Schnittstelle anhand von Heuristiken durch diverse Experten untersucht, indem die Schnittstelle im Hinblick auf Probleme bzw. Gestaltungsmängel, aber auch auf Vorteile betrachtet wird.

Die leitfadenbasierte Evaluation erfolgt hingegen nach vorher definierten Ablaufprozessen, so dass nur diejenigen Fehler durch Experten gefunden werden können, die zuvor auch als möglich betrachtet wurden.

Beim Cognitive Walkthrough (CW) werden vorgegebene Handlungsabläufe von Experten mit einem System durchgeführt und hinsichtlich ihrer Ausführbarkeit untersucht. Bei der Durchführung eines CWs stehen die mentalen Modelle der Nutzer im Fokus. Ziel ist es

Produkte bezüglich ihrer intuitiven Benutzbarkeit und Erlernbarkeit zu evaluieren (Heuer, 2003).

Bei den nutzerbasierten Evaluationsverfahren werden Anwender hinzugezogen, um Systeme zu untersuchen. Nutzerbasierte Verfahren werden oft mit dem Oberbegriff "User-Test" bezeichnet. Dabei kann es sowohl um verschiedene Untersuchungsschwerpunkte wie Usability oder Nutzererleben gehen, als auch um verschiedene Methodenkombinationen. Neben dem Interviewverfahren und den objektiven Messungen gibt es bei den User-Tests noch das Fragebogenverfahren und die Methode des Lauten Denkens.

Bei der Methode des Lauten Denkens soll die Testperson während der Interaktion mit dem Produkt ihre dabei auftretenden Gedanken laut äußern. Dies ermöglicht Einblicke in das mentale Modell der Nutzer. Mit Hilfe dieser Methode können Rückschlüsse auf Eindrücke, Gefühle und Intentionen gezogen werden. Diese Methode ist eine häufig eingesetzte Methode, da hier der Nutzer eine Aufgabe mit dem System unter Verwendung des Lauten Denkens ausführt und Probleme sofort sichtbar werden. Allerdings ist die Auswertung der Aussagen, aufgrund der erforderlichen Transkription sehr aufwändig.

Fragebogenverfahren hingegen werden vorzugsweise angewandt, um beispielsweise die Nutzerzufriedenheit in Bezug auf die Qualität (z.B. mit dem AttrakDiff) oder die mentale Beanspruchung des Nutzers durch das System zu untersuchen (wie z.B. beim NASA-TLX).

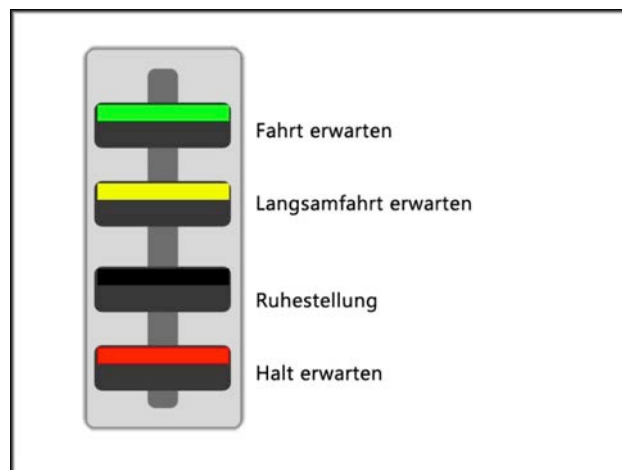
Für die Evaluation der Interfaces im Rahmen der Projektarbeit wurden der NASA-TLX, der ISONORM 9241/110-S, die Methode des Lauten Denkens und ein abschließendes Interview verwendet. Auf diese Methoden wird in den folgenden Kapiteln noch genauer eingegangen. Die Materialien sind im Appendix nachzulesen.

## 4 Konzeption und prototypische Umsetzung

### 4.1 Hebel

Das entwickelte Bedienelement basiert auf dem Konzept eines Hebels, der eingleisig nach vorne und zurück bewegt werden kann. Durch ein Hebelmoment ist der Benutzer in der Lage, mechanisch durch Drücken beziehungsweise Ziehen zwischen den Stellungen hin und her zu schalten und den Hebel so in die gewünschte Position zu bringen. Dabei ist die

Eingabe der Vorsignale “Fahrt erwarten”, “Langsamfahrt erwarten” sowie “Halt erwarten” möglich. Zusätzlich gibt es die Position “Ruhestellung”. Wird beispielsweise das Vorsignal “Fahrt erwarten” vom Lokführer erkannt, so betätigt er den Hebel von der Ruhestellung durch Schieben zwei Positionen nach oben. Der Hebel liegt nun in “Fahrt erwarten”. Erst bei der Quittierung des nächsten Vorsignals soll der Hebel in die Ruhestellung gebracht werden, um im Anschluss direkt das neue Vorsignal zu quittieren. Dieses Vorgehen hat den Vorteil, dass dem Lokführer auf der Strecke von Vorsignal 1 zu Vorsignal 2 ein Feedback über die Hebelstellung ermöglicht wird.



**Abbildung 2: Hebel-Konzept**

Die Umsetzung des Konzeptes basiert auf der Verwendung von Analogien. Beispielsweise wird auch in der Luftfahrt ein Schubhebel für die Geschwindigkeitssteuerung verwendet. Das Bedienelement soll eine kognitive Anreicherung erzwingen, da das Vorsignal inhaltlich interpretiert werden muss, um die jeweilige Einstellung vornehmen zu können. Damit soll die bewusste Verarbeitung des Vorsignals gewährleistet werden. Nach der Quittierung bietet die verbleibende Hebelstellung zusätzliche Informationen über die aktuelle Vorsignal-Einstellung.

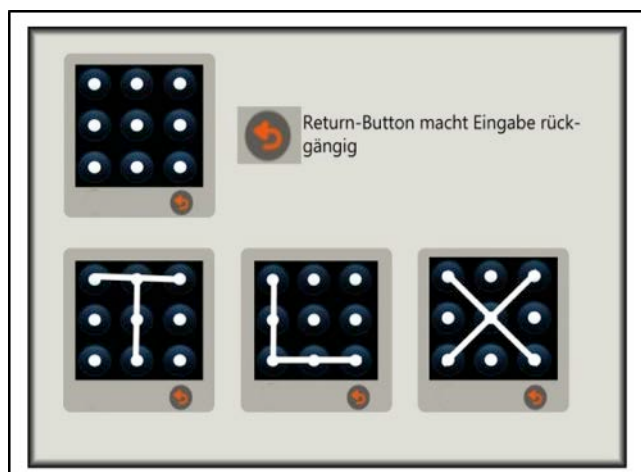
Technisch umgesetzt wurde dieses Konzept durch ein Schaltungselement, welches sich aus einem Drehschalter DS 2x12LF/B mit vier Stellungen und einer 6mm-Achse zusammensetzt, auf dem der Hebelarm, bestehend aus Metallplatte und Knauf über eine Federspannung befestigt wird. Für die Anzeige werden acht LEDs in den Farben grün, gelb, blau und rot verwendet. Diese werden über eine externe Stromquelle bestehend aus drei 1.5V Batterien gespeist. Vier LEDs werden im Metallgehäuse und vier auf der Platine angebracht, die dem Versuchsleiter die gewählten Anzeigen darstellen.



**Abbildung 3: Hebel-Prototyp**

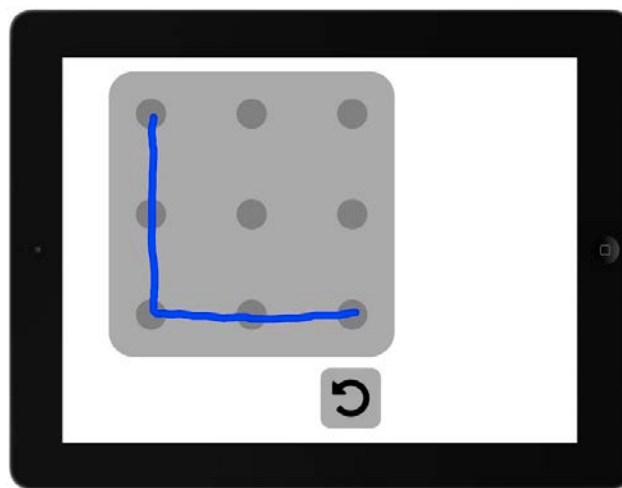
## 4.2 Touchscreen

Das zweite Konzept basiert auf einem Touchscreen, welches die Bedienung per Finger ermöglicht. Hierfür wurde eine Zeichen-App erstellt, die im Hintergrund eine 3x3-Punktvorlage besitzt. Dieser Hintergrund erleichtert dem Benutzer das Zeichnen der Symbole, da eine bessere Orientierung im Bild erfolgt. Jedes Symbol besteht aus zwei Strichen, weshalb sie vom zeitlichen Aufwand vergleichbar sind. Auch die räumliche Auslastung ist vergleichbar, da immer die Außenpunkte benutzt werden.



**Abbildung 4: Touch-Interface-Konzept**

Die unterschiedlichen Vorsignale werden durch das Zeichnen von drei Buchstaben quittiert. Diese drei Zeichen beschreiben jeweils die passenden Vorsignale. Das Symbol T wird für das Vorsignal “freie Fahrt” verwendet. “Langsamfahrt erwarten” wird durch den Buchstaben L verdeutlicht und das Symbol X beschreibt “halt erwarten”. Diese Symbole sind nicht frei gewählt, sondern verdeutlichen Metaphern. Die Übersetzung des Vorsignals in die abstrakteren Symbole X, L und T bedeutet einen höheren kognitiven Aufwand. Die Symbole sind nicht durch eigene Assoziationen vorbelastet, aber können im Training leicht mit deren Bedeutung in Verbindung gebracht werden. Der Buchstabe T soll das Wort “Tempo” verdeutlichen, L steht für “langsam” und X beschreibt “halt”.



**Abbildung 5: Touch-Interface-Prototyp**

Falls der Benutzer eine falsche Eingabe tätigt, besteht die Möglichkeit das zuvor gezeichnete zu korrigieren. Dafür existiert ein Return-Button unter dem Zeichenfeld. Nachdem das Symbol gezeichnet wurde bleibt es für 4 Sekunden sichtbar. Dadurch, dass es sichtbar bleibt, steht dem Lokführer das Symbol zur Verfügung falls vergessen wird was gezeichnet wurde. Es kann folglich eine erneute Wahrnehmung des Vorsignals unterstützen. Nach den 4 Sekunden verschwindet das Symbol, denn das Zeichenfeld muss für weitere Vorsignale bereit sein.

Die Quittierung erfolgt in dem Moment, wo das Vorsignal passiert wird. Zusätzlich kann nach den 4 Sekunden eine Verarbeitung der Eingabe durch die PZB erfolgen. So bleibt die Kontrolle des Lokführers erhalten.



## 5 Methode

### 5.1 Stichprobe

#### 5.1.1 Rekrutierung

Für die Rekrutierung der Probanden wurden verschiedene Wege beschritten. Via E-Mail, Telefon und den Direktkontakt am Institut für Land- und Seeverkehr der Abteilung Schienenfahrwege und Bahnbetrieb der TU Berlin wurde Werbung für den Versuch gemacht. Für den elektronischen Weg wurde ein Flyer erarbeitet, der Interesse am Projekt wecken sollte und die wichtigsten Informationen enthält. Unter anderem wurden die DB, BVG, ODEG und verschiedenste Ehrenlokführerseiten adressiert. Außerdem wurde über Facebook, Twitter und sofern möglich, persönliche Kontakte auf den Versuch aufmerksam gemacht. Beginnend im Dezember 2013 verlief der Rekrutierungsphase auf Grund mangelnder Probanden bis Mitte Januar 2014. Letztlich konnten vier Probanden für den Versuch gewonnen werden.

#### 5.1.2 Beschreibung der Stichprobe

Die Stichprobe setzt sich aus vier männlichen Lokführern zusammen. Darunter waren zwei jüngere Teilnehmer (25; 26) und zwei ältere Teilnehmer (42; 48). Alle Versuchspersonen haben bereits Erfahrung im Umgang mit Touch-Displays und nehmen regelmäßig an Weiterbildungen teil. Die Berufserfahrung der Probanden reicht von einem Jahr bis zu 28 Jahren. Drei der vier Teilnehmer sind im Güterverkehr tätig. Zwei Probanden fahren hauptsächlich Regionalbahn. Die Hälfte der Teilnehmer ist auch außerhalb von Deutschland, vorrangig in der Schweiz, Österreich und den Niederlanden tätig.

### 5.2 Versuchsdesign

Für das Versuchsdesign wurde sich stark am User Centered Design orientiert. Daher stand der Endnutzer in diesem Fall der Lokführer, im Mittelpunkt unseres Versuchs. Über verschiedene quantitative und qualitative Ansätze, die im Folgenden noch genauer beschrieben werden, wurden die Eindrücke, Bedürfnisse und Erfahrungen des Nutzers mit den Interfaces erfasst. Es wurde eine formative Evaluation durchgeführt, da die Meinung des Nutzers in die weitere Konzeption und Ausarbeitung der Interfaces für mögliche Folgeprojekte mit einfließen soll.

## 5.2.1 Materialien

### 5.2.1.1 *Lautes Denken*

Während des Versuchs wurde die Technik des Lauten Denkens angewendet. Dabei wird der Proband dazu angehalten, Gedanken direkt und frei zu äußern. Daher ist es möglich, konkrete und vor allem spontane Eindrücke des Probanden während der Interaktion mit den unterschiedlichen Interfaces zu gewinnen, ohne, dass der Eindruck durch Erinnerungsverzerrungen, sowohl seitens des Probanden als auch des Versuchsleiters, verfälscht wurde. Für die Datenauswertung wird das Laute Denken per Tonband aufgenommen. Der Ablauf (*Anhang C*) und der Instruktionstest zum lautem Denken sind im Appendix aufgeführt (*Anhang D*).

### 5.2.1.2 *NASA-TLX*

Der NASA-TLX ist ein subjektives, mehrdimensionales Befragungsmaß, welches die Beanspruchung der Probanden erfasst. Die Gesamtbeanspruchung wird dabei über sechs Subskalen erfasst. Diese sind: Geistige Anforderungen, Körperliche Anforderungen, Zeitliche Anforderungen, Ausführung der Aufgaben, Anstrengung und Frustration. Für jede Subskala liegt dem Probanden eine Beschreibung der jeweiligen Anforderung vor. Das Antwortformat umfasst eine 11-stufige Ratingskala von “gering” bis “hoch” pro Subskala. Jede Subskala liefert also ein Item. Daher ist der Fragebogen mit sechs Items sehr ökonomisch. Im Anhang befindet sich die hier verwendete deutsche Übersetzung des Fragebogens (*Anhang F*).

### 5.2.1.3 *ISONORM 9241/110-S*

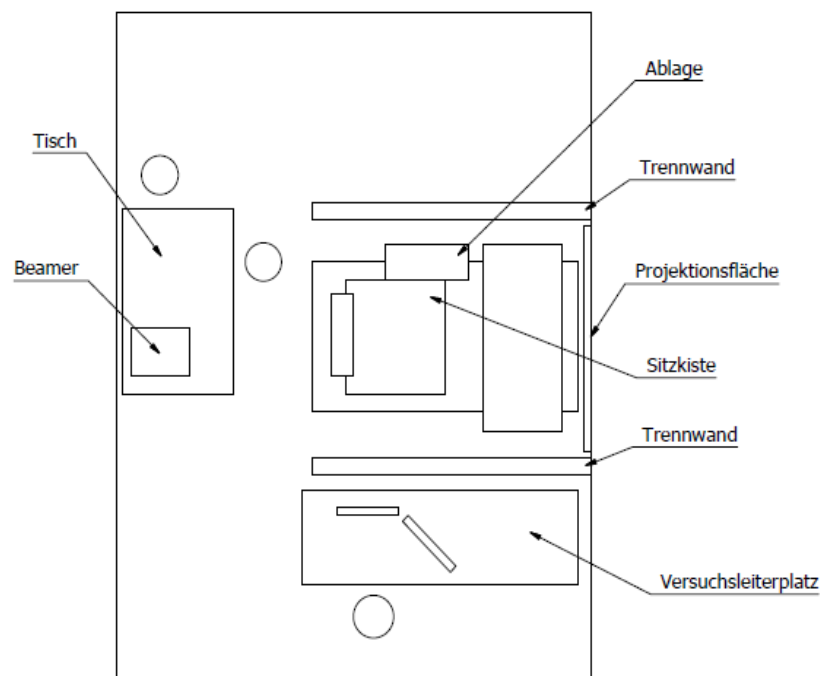
Der ISONORM 9241/110 ist ein Fragebogen, der die Beurteilung der Usability von Software erfasst. Die hier verwendete Kurzform umfasst 21 der erstmals 35 Items. Grundlage für den Fragebogen ist die DIN EN ISO 9241, die in ihrem zehnten Teil die Grundsätze der Dialoggestaltung aufführt. Diese sind: Aufgabenangemessenheit, Selbstbeschreibungsfähigkeit, Steuerbarkeit, Erwartungskonformität, Fehlertoleranz, Individualisierbarkeit und Lernförderlichkeit. Jeder Grundsatz wurde im Fragebogen durch eine Subskala operationalisiert. Jede Subskala umfasst drei Items. Das Antwortformat umfasst eine siebenstufige Likertskala von sehr negativ (---) bis sehr positiv (+++). Darüber hinaus lassen sich mit dem Fragebogen Aussagen über Schwachstellen der verwendeten Interfaces machen. Für eine bessere Passung des Fragebogens mit dem vorliegenden Versuchsdesign wurde das

Wort “Software” mit dem Wort “Interface” ausgetauscht. Für genauere Informationen über die Items findet sich auch dieser Fragebogen im *Anhang G*.

#### 5.2.1.4 Interview

Abschließend wurde ein etwa 10-minütiges Interview auf Basis eines halbstandardisierten Interviewleitfadens durchgeführt. Dieser hatte den Zweck, dem Probanden die Möglichkeit für ein abschließendes Urteil bezüglich der zwei getesteten Interfaces zu geben. Außerdem wurde ein Vergleich der Interfaces miteinander sowie der Vergleich der Interfaces mit der derzeitigen Quittierungshandlung thematisiert. Darüber hinaus diente das Interview als erweitertes Validierungsinstrument, da die Interviewdaten mit den Äußerungen, die im Rahmen des Lauten Denkens erfasst wurden, auf Übereinstimmung geprüft werden konnten. Der Interviewleitfaden ist im Anhang aufgeführt (*Anhang H*).

### 5.2.2 Versuchsaufbau



**Abbildung 6: Versuchsaufbau**

Der Versuchsaufbau (*Abb. 6*) wurde in einem Raum des Marchgebäudes der TU Berlin in einer ruhigen Umgebung umgesetzt. Dies ermöglicht konstante Verhältnisse, so dass alle störenden audiovisuellen Reize vermieden werden. Während der Tests wurden die Fenster durch Rollos verdunkelt und die Zimmerbeleuchtung ausgeschaltet. Im Untersuchungsraum befanden sich während des Versuchs nur der Proband, ein Versuchsleiter und ein

Techniker. Der Versuchsaufbau besteht aus einer Sitzkiste, in dem der Proband Platz nahm. Auf der linken Seite der Sitzkiste befand sich eine Ablage, auf der die Konzepte je nach Versuch angebracht wurden.

Die Streckenvideos (Auflösung 1280x720, Dauer 20:40 (9 Vorsignale) und 12:42 (10 Vorsignale), Videoabschnitte von DB Führerstandmitfahrt №23 (Teil 2) und №1, beide verfügbar auf YouTube) wurden mit Hilfe eines Beamers auf die Fläche vor dem Probanden projiziert und für den Versuchsleiter auf dem Versuchsleiterplatz auf einem von zwei Bildschirmen geklont. Der zweite Bildschirm am Versuchsleiterplatz diente der gespiegelten Wiedergabe von Eingaben auf dem iPad.

Um eine mögliche Ablenkung durch den Versuchsleiter oder den Techniker zu verhindern, wurden links und rechts von der Sitzkiste Trennwände aufgestellt. Zur Aufzeichnung des lauten Denkens während des Versuches diente ein Diktiergerät, das vor dem Probanden positioniert wurde. Für die Auswertung der Konzepte wurden die Eingaben der Probanden mittels einer Webcam und dem Aufzeichnungstool “*VideoDub*” erfasst. Für die Bearbeitung der Fragebögen und das anschließende Interview wurde ein Arbeitsplatz hinter der Sitzkiste im Labor zur Verfügung gestellt.

### 5.2.3 Versuchsaufbau

Der Versuch umfasste etwa eineinhalb Stunden. Nach einer Begrüßung und einigen einführenden Worten wurde ein Informationsblatt ausgehändigt, was alle wesentlichen Informationen über den Versuch, Datenschutz, sowie über den Zeitumfang und die Vergütung des Versuchs beinhaltete (*Anhang B*). Nachdem bezüglich des Versuchs und der Tonaufnahme für das Laute Denken das Einverständnis schriftlich eingeholt wurde, wurden einige demografische Daten per Fragebogen erhoben (*Anhang A*). Anschließend konnte der Proband in der Sitzkiste Platz nehmen, wo via Beamer die Fahrtstrecke zu sehen war. Vor Ort wurde die Versuchsperson gebeten, ein Informationsblatt zum Lauten Denken laut vorzulesen (*Anhang D*). Diese Übung sollte die Versuchsperson noch einmal über die Hintergründe des Lauten Denkens informieren und auf die Aufgabe vorbereiten. Das erste Interface wurde vorgestellt und ein Informationsblatt über die Bedienung wurde ausgehändigt. Mit einer kleinen Übung, bei der sechs Vorsignale dargeboten wurden, die mittels Interface richtig quittiert werden sollten, wurde der Umgang mit dem Interface erlernt. Etwaige Fragen, die sich aus der Übung ergaben, konnten direkt geklärt werden. Im Anschluss wurde die erste Fahrtstrecke dargeboten. Folglich setzte sich der Proband wieder an den Tisch,

wo ihm vom Versuchsleiter zuerst der NASA-TLX und anschließend der ISONORM 9241/110-S ausgehändigt wurden. Für das zweite Interface wurde der Ablauf wie beschrieben mit einer anderen Fahrstrecke wiederholt. Die Interfaces (Hebel; Touchscreen) sowie die Fahrten (Fahrt 1, ca. 21 Minuten; Fahrt 2, ca. 11 Minuten) wurden randomisiert dargeboten, um Reihenfolgeeffekte zu vermeiden.

Nachdem auch die zweite Fahrt absolviert wurde, führte der Versuchsleiter das Interview mit dem Probanden durch. Abschließend wurde die Teilnahme schriftlich bestätigt, die Vergütung von 20 € ausbezahlt und der Proband verabschiedet.

### 5.3 Auswertungsmethoden

Für die Auswertung des Lauten Denkens werden die Daten zuerst via Tonbandaufnahme transkribiert. Anschließend erfolgte die *Qualitative Inhaltsanalyse nach Mayring* (2010). Qualitative Methoden ermöglichen in Abgrenzung zu quantitativen Methoden ein besseres Verständnis von den Motiven der handelnden Person sowie gegebenenfalls die Einbeziehung der Gedanken und Gefühle der Person. Um dem User Centered Design gerecht zu werden, kommt der Verwendung qualitativer Erhebungsmethoden daher eine besondere Bedeutung zu. Die Qualitative Inhaltsanalyse wurde mit dem Ziel der induktiven Kategorienbildung angewandt. Dabei wird das vorliegende Textmaterial im Hinblick auf ein Selektionskriterium, welches den Ausgangspunkt der Analyse festlegt, durchgegangen. Ausgangspunkt der Analyse sind in dieser Studie Textpassagen, die sich inhaltlich auf die Bedienbarkeit der Interfaces beziehen. Inhaltlich passende Textpassagen werden zusammengefasst und schrittweise verallgemeinert (Generalisierung), bis sie letztlich zu einer Kategorie zusammengetragen werden können (Reduktion). Im Anschluss wird das Kategoriensystem anhand der Transkription noch einmal rücküberprüft, um sicherzustellen, dass die Kategorien das Ausgangsmaterial noch hinreichend abbilden. Bei der induktiven Kategorienbildung werden also Kategorien aus dem Textmaterial erschlossen.

Zur Auswertung des qualitativen Interviews wurden zunächst die protokollierten Aussagen der Probanden den jeweiligen Interface-Typen zugeordnet. Bei der Auswertung sollte eine grobe Orientierung an den thematischen Schwerpunkten des Interviewleitfadens vorgenommen werden. Deshalb wurden zunächst mithilfe der erfragten Themenbereiche die Kategorien Bedienbarkeit, Sicherheit, Assoziationen, Erlernbarkeit, Probleme und Verbesserungsvorschläge abgeleitet. Danach konnten die Aussagen den jeweiligen Kategorien

zugeordnet werden. Dabei war es möglich, dass eine Aussage verschiedenen Themengebieten zugeteilt wurde. Im letzten Schritt wurden die Antworten innerhalb der jeweiligen Kategorien zusammengefasst.

Sowohl die Daten der Fragebögen NASA-TLX und ISONORM 9241/110-S als auch die Fehlerrate wurden deskriptiv analysiert. Darüber hinaus wurde der Wilcoxon-Test für Paardifferenzen bei einem Signifikanzniveau  $\alpha = 0.05$  angewendet. Die Gründe für die Auswahl dieses Tests sind folgende: dieselben Teilnehmer haben die beide Interfaces überprüft (verbundene Stichprobe), die Stichprobe ist sehr klein ( $n = 4$ ) und beim Wilcoxon-Test handelt es sich um eine verteilungsfreie Alternative zum t-Test (Bortz & Schuster, 2010). Der t-Test hingegen setzt für kleine Stichproben eine Normalverteilung voraus. Wenn eine Abweichung von der Normalverteilung vorliegt, sollte ein nichtparametrisches Verfahren angewendet werden (Fahrmeir, Künstler, Pigeot, & Tutz, 2007). Aus diesem Grund wurde dieser für die Auswertung angewendet. Es ist zu berücksichtigen, dass die inferenzstatistische Ergebnisse kritisch betrachtet werden sollten, da man anhand der extrem kleinen Stichprobe nur sehr vorsichtig und ziemlich unsicher Aussagen treffen kann.

Für statistische Analyse wurde die Statistiksoftware *IBM SPSS Statistics 22* eingesetzt.

Da eine Frage aus der Dimension "Fehlertoleranz" in ISONORM 9241/110-S ausgeschlossen wurde, müsste man diese Dimension gewichten, damit sie mit den anderen Subskalen vergleichbar ist.

## 6 Ergebnisse

### 6.1 Lautes Denken

Aus den thematischen Schwerpunkten der Verbalisierungen während des Umgangs mit den beiden Interfaces wurden induktiv verschiedene Kategorien abgeleitet. Bei dem Hebel waren dies die Kategorien (a) informative Wirkung: umfasst Aussagen zum Feedback des Hebels, (b) Bedienung: enthält alle Äußerungen bezüglich des Umgangs während der Quittingshandlung, (c) Hebel irreführend: bezieht sich auf Eigenschaften des Hebels, die von den Versuchspersonen als unverständlich oder verwirrend beschrieben wurden, (d) Ablenkung: Aussagen, mit der die Bedienung bezüglich seiner Ablenkung eingeschätzt

wurde, und (e) Spaß: Aspekte, die von den Probanden als positiv, angenehm und Spaßig beschrieben wurden. Für das Touch-Interface wurden insgesamt drei Kategorien gebildet. Diese sind die Kategorien (f) Feedback: Aussagen, bei denen die fehlenden Rückmeldungen über die Eingaben thematisiert wurden, (g) Bedienung zu ablenkend: Bewertungen, bei denen der Touchscreen als aufwändig und störend beschrieben wurde und (h) Intuition: Aussagen bezüglich der Verständlichkeit der Bedienung.

Insgesamt wurde der Hebel von allen Versuchspersonen favorisiert und dem Touch-Interface vorgezogen. Ein wichtiger Grund dafür war die bessere Bedienbarkeit des Hebels aufgrund seiner leicht zu erlernenden und intuitiv verständlichen Quittierungshandlung. Verglichen mit dem Touchpad ermöglichte der Hebel eine blinde Bedienung, was eine konstante Beobachtung der Strecke gewährleistete. Eine Versuchsperson merkte an, dass sich das Hebel-Interface aufgrund seiner einfachen Bedienung auch für Auszubildende gut eigne („[...] mit der Bedienung kommen bestimmt auch Neulinge gut zurecht, da die Ausbildung schon sehr kompakt ist und stressig.“, „Es geht fast von selbst. Erlernt in wenigen Minuten eigentlich.“). Das Touchpad hingegen benötigte sehr viel Aufmerksamkeit und Konzentration.

Die Probanden mussten den Blick oft von der Strecke auf das den Touchscreen richten, um die jeweiligen Symbole korrekt einzuzeichnen. Alle Probanden bemängelten dabei, dass diese Form der Quittierung zu viel Zeit in Anspruch nehme („Auf geraden Strecken kein Problem, da kann man rechtzeitig Fahrt wegnehmen, aber bei Kurven [...]“, „In der Praxis wird's schwierig werden, weil man gleichzeitig Geschwindigkeit reduzieren muss. Man braucht beide Hände.“, „Ich stell mir das auch schwer vor, wenn ich da 160 fahre oder 140, jedes Mal auf das Display zu gucken, ob ich das richtig gemacht habe oder nicht.“). Einen weiteren Kritikpunkt betraf die Notwendigkeit das Symbol innerhalb des vorgegeben Rasters einzuzeichnen. Dies würde einen zusätzlichen Konzentrationsaufwand verursachen, der sich negativ auf die eigentliche Fahraufgabe auswirke („Hier hätte ich Angst, dass ich das richtige Feld nicht treffe und dann ne Zwangsbremse kriege.“). Der Hebel hingegen konnte von allen Probanden schnell verwendet werden, ohne dass sie dabei ihren Blick nach draußen unterbrechen mussten („Blick kann länger auf Straße bleiben.“, „Klappt sofort ohne hinzugucken.“).

Eine Versuchsperson empfand den Umgang mit dem Hebel sogar als Spaßig („[...] Macht an sich aber richtig Spaß. Hat das Lokfahren ein ganz anderes Feeling.“). Insbesondere

im Vergleich zum Touch-Interface, bei dem die Eingabe unter Einsatz starker visueller Aufmerksamkeit getätigt werden muss, könne man mit dem Hebel-Interface eine entspannte Haltung einnehmen und sich besser auf das Fahren konzentrieren (*„Ich muss nicht starr sitzen. Ich kann entspannt sitzen und dabei den Hebel in Bewegung setzen. Kann mich also in Ruhe auf das Vorsignal konzentrieren, ohne darauf zu achten, ob ich jetzt ein T, X oder ein L ins iPad zeichne.“*).

Es wurde außerdem als positiv erachtet, dass die jeweiligen Hebelstellungen und das Aufleuchten der dazugehörigen LED-Lampen stets als Referenz über das zu erwartende Signal verwendet werden können. Allerdings schienen die LEDs, insbesondere auf langen Strecken ohne Signalmeldungen, einen gewissen Aufforderungscharakter zu haben (z.B. *„(rote) LED symbolisiert, da droht noch Gefahr!“*). Für beide Interfaces wurde die Integration eines Feedbacks über die getätigten Signaleinstellungen vorgeschlagen. Der Hebel würde sich beispielsweise besser bedienen lassen, wenn eine Einrastfunktion vorhanden wäre. Und auch das Touch-Interface wurde aufgrund seiner mangelnden Rückmeldung zur Bestätigung der eingezeichneten Symbole kritisiert (*„Ohne Feedback ist doof.“*, *„[...]als wenn man ins Leere tippt.“*). Hier schlugen viele Versuchspersonen ein haptisches oder optisches Feedback vor.

## 6.2 Interview

Um die Meinungen der Versuchspersonen zu erfragen, wurde im Anschluss des Versuchs ein Interview durchgeführt. Dabei wurden die Probanden darum gebeten die beiden Interfaces hinsichtlich ihrer Bedienbarkeit und Erlernbarkeit zu vergleichen. Auch sollten sie ein Vergleich zum aktuellen Quittierungssystem vornehmen. Außerdem wurde nach Problemen und möglichen Verbesserungsvorschlägen gefragt.

Insgesamt fielen die Meinungen über die Interfaces bei allen Probanden ähnlich aus. Das Hebel-Konzept wurde von allen Probanden am besten bewertet und akzeptiert. Den Touchscreen hingegen empfanden sie als gewöhnungsbedürftig. Dabei wurde insbesondere die verstärkt aufzubringende Aufmerksamkeit und Konzentration kritisiert. Von dem Hebel-Interface hingegen wurde angenommen, dass er mit ein wenig Übung blind bedient werden könnte. Dies wurde von allen Versuchspersonen als wichtiger Aspekt zur Gewährleistung der Sicherheit betont, weil so eine stete Beobachtung der Strecke sichergestellt ist.



Der Quittierungshandlung mit dem Touch-Interface standen die Probanden eher skeptisch gegenüber. Sie wurde als sehr komplex und zeitaufwändig empfunden. Die Übersetzung der Vorsignale in die entsprechenden Symbole beanspruchte die Probanden. Die zusätzliche Notwendigkeit die Buchstaben innerhalb eines Rasters einzuzeichnen, wurde von den Probanden als “schwer” und “zu viel” empfunden. Ein Vermalen würde außerdem zusätzliche Tätigkeiten umfassen, sodass auch hier die Ablenkung viel zu groß sei.

Mit dem Hebel-Interface würden sich die Probanden sicherer fühlen. Es wurde insbesondere die gute Haptik und die gebräuchliche Handhabung als positiv erachtet. Der Aufwand der Quittierungshandlung wurde als angemessen eingeschätzt. Auch die Bedienungslogik wurde als sinnvoll erachtet, da das Vor- und Zurückschieben des Hebels analog zu anderen Systemen funktioniere. Die zusätzliche Beschriftung der jeweiligen Hebelstellungen und die LED-Beleuchtungen wurden als Erleichterung für die Bedienung erkannt. Positiv wurden auch die bestehenden Hebelstellungen als Erinnerungstütze über das aktuelle Signal bewertet.

Kritisiert wurde hingegen die Ruhestellung. Die Notwendigkeit den Hebel am Vorsignal zunächst in die Ruhestellung zu bringen bevor das neue Signal quittiert wird, sorgte für Verwirrung. Es wurde stattdessen vorgeschlagen, die Ruhestellung entweder direkt nach der Quittierung des Vorsignals vorzunehmen oder ganz wegzulassen. Es wurde außerdem darauf hingewiesen, dass sich bereits viele Hebel und Schalter im Cockpit befinden, was eine optische Abgrenzung eventuell erschweren könnte.

Für beide Interfaces wurden Feedback-Mechanismen zur Rückmeldung der getätigten Eingaben, wie zum Beispiel Vibration beim Touchscreen oder einer Einrastfunktion beim Hebel, vorgeschlagen.

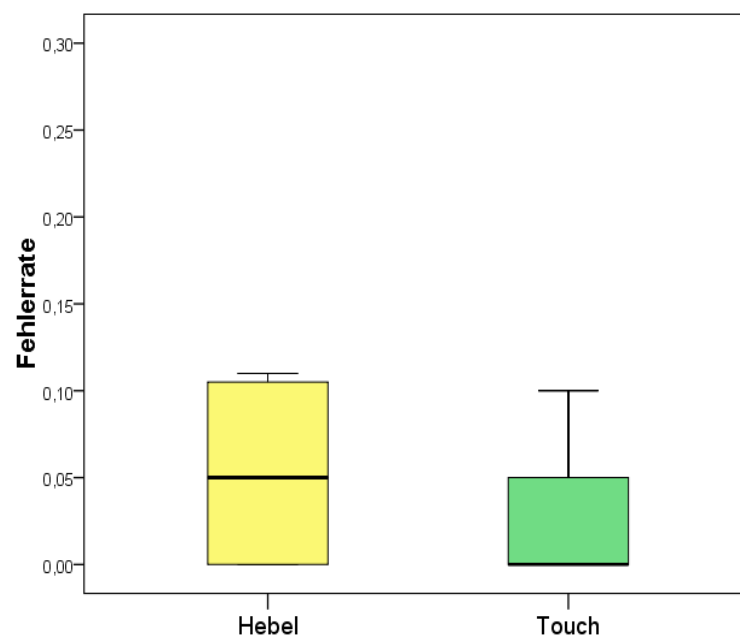
Keiner der Probanden hatte das Gefühl, durch eines der beiden Interfaces die Vorsignale bewusster zu verarbeiten. Allerdings hielt eine Versuchsperson die Farbkodierung beim Hebel-Konzept für gut geeignet, um sich die jeweiligen Vorsignal-Einstellungen merken zu können. Die Buchstaben beim Touchscreen hingegen seien nur schwer mit den jeweiligen Vorsignalen vereinbar und wurden als sehr zeitaufwendig empfunden.

Es zeichnete sich zudem ab, dass die Probanden lieber am bestehenden System festhalten wollen. Die aktuelle Quittierungshandlung mit dem Wachsamkeitsschalter sei routiniert

und ließe sich einfach ausführen. Eine Versuchsperson merkte hierzu an, dass es schwierig sei Lokführer umzulernen.

### 6.3 Fehlerrate

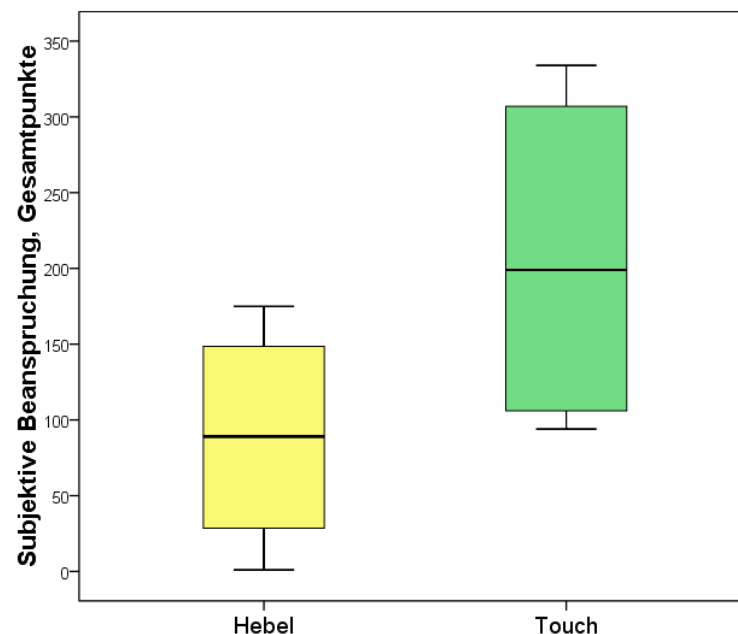
Als objektives Maß wurde die Fehlerrate analysiert. Die Fehlerraten wurden als Verhältnis von aufgetretenen Fehlern zu allen relevanten Ereignissen (Vorsignale) definiert und berechnet. Als Fehler wurde die falsche Einstellung des Hebels bzw. eine falsche Eingabe auf dem Touchscreen bezeichnet. Ob die Fehler jedoch tatsächlich durch die Interfaces verursacht wurden, lässt sich nicht mit Sicherheit feststellen. Daher sind weitere Untersuchungen der Fehleranfälligkeit der Interfaces empfehlenswert. Allgemein liegen die Fehlerwerte im unteren Bereich: die maximale Fehlerrate beträgt 0.11 (Abb. 7). Die Verteilung der Fehlerrate in der Stichprobe unter der Hebel-Bedingung ist symmetrisch, die Verteilung unter der Touchscreen-Bedingung ist stark linksseitig. Der Wilcoxon-Test für Paardifferenzen hat keinen signifikanten Unterschied ( $\alpha = 0.05$ ) in der Fehlerrate zwischen dem Hebel ( $MD = 0.05$ ) und dem Touchscreen ( $MD = 0$ ) entdeckt,  $z = -1.342$ ,  $p = 0.180$ .



**Abbildung 7: Boxplot-Diagramm für die Verteilung der Fehlerrate bei der Nutzung des Hebels und des Touchscreens**

## 6.4 Subjektive Beanspruchung (NASA-TLX)

Das Boxplot-Diagramm für die Verteilung des Gesamtwertes zur subjektiven Beanspruchung bei der Benutzung der Interfaces zeigt, dass der Hebel von den Versuchsteilnehmern insgesamt weniger beanspruchend wahrgenommen wurde als das Touch-Interface. Deskriptiv lassen sich Unterschiede statistischer Parameter, wie das untere und oberes Quartil und der Median, feststellen. Auffällig ist, dass der höchste Wert der Gesamtbeanspruchung für das Hebel-Interface unter dem Median des Touchscreens liegt. Die beiden Verteilungen sind annähernd symmetrisch, aufgrund dessen eine Normalverteilung angenommen wird. Für die Bewertung des Touch-Interfaces ist eine größere Streuung zu beobachten. Inferenzstatistisch wurde der Unterschied zwischen dem Hebel und dem Touchscreen allerdings nicht signifikant:  $z = 1.826$ ,  $p = 0.068$ .



**Abbildung 8: Boxplot-Diagramm für die Verteilung des gesamten Wertes zur subjektiven Beanspruchung bei der Nutzung des Hebels und Touchscreens**

Die dem Boxplot-Diagramm (Abb. 9) zu entnehmenden Bewertungen auf den sechs Subskalen der subjektiven Beanspruchung liefern Hinweise auf tendenziell unterschiedlich erlebte Beanspruchungen mit den beiden Interfaces: der Hebel wurde in fast allen Subskalen, bis auf die der körperlichen Anforderungen, als wenig beanspruchend bewertet. Es fällt auf, dass große individuelle Unterschiede zwischen den Bewertungen vorliegen. Dies ist insbesondere beim Touchscreen der Fall, bei dem die Streuung der Werte für die Fakto-

ren Anstrengung, Aufgabenausführung, Frustration sowie geistige und zeitliche Anforderung stärker ausgeprägt ist.

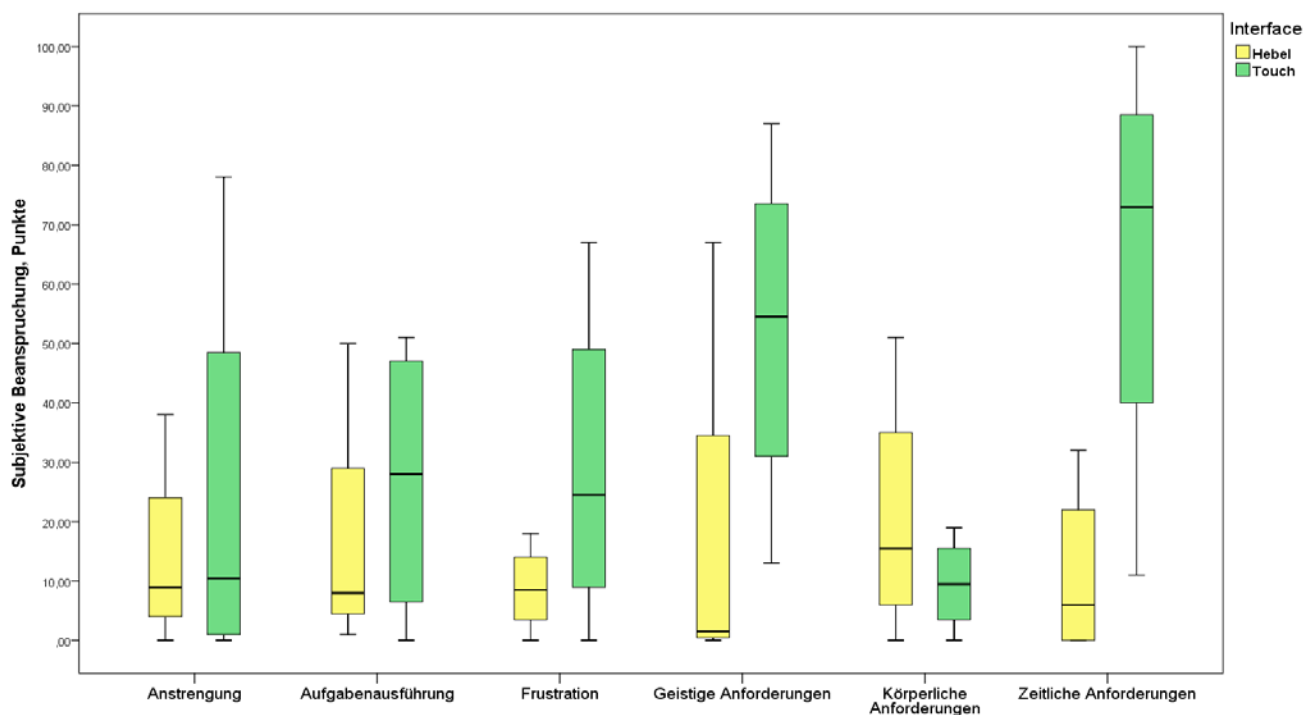
Besonders die Verteilungen der Skalen *geistige Anforderung* und *zeitliche Anforderung* weichen stark von den übrigen Subskalen ab. Es ist zu erkennen, dass das Touch-Interface tendenziell als zeitaufwendiger und kognitiv beanspruchender erlebt wurde als der Hebel. Die Bewertungen der subjektiven Beanspruchung fallen insgesamt geringer für das Hebel-Interface aus als für den Touchscreen (deutliche Unterschiede bei den unteren und oberen Quartilen sowie den Medianen) und verteilen sich stark linkslastig: sowohl die Mediane der geistigen Anforderung, als auch die des Faktors zeitliche Anforderung liegen im unteren Wertebereich. Im Gegensatz dazu sind die Verteilungen der Bewertungen des Touch-Interfaces leicht rechtslastig und enthalten deutlich höhere Werte. Obwohl weder die Median-Differenzen bezüglich der Dimension der geistigen Anforderung ( $MD$  (Hebel) = 1.5,  $MD$  (Touch) = 54.5) noch zu der zeitlichen Anforderung ( $MD$  (Hebel) = 6,  $MD$  (Touch) = 73) mit dem Wilcoxon-Test auf dem Signifikanzniveau  $\alpha = 0.05$  signifikant wurden ( $z = 1.826$ ,  $p = 0.068$ ;  $z = 1.461$ ,  $p = 0.144$ ), liegt zumindest deskriptiv die Tendenz eines Unterschiedes zwischen den Interfaces vor.

Weitere Unterschiede zwischen den Interfaces sind für die Subskala *Frustration* zu finden. Die Verteilungen der Bewertungen unterscheiden sich bezüglich ihrer Streuung: die Werte des Touch-Interfaces liegen weit auseinander und die Verteilung ist leicht linkslastig, während die Bewertungen zum Hebel eine symmetrische Verteilung mit geringerer Streuung darstellen. Bemerkenswert dabei ist, dass die höchste Bewertung für die Skala Frustration beim Hebel unter dem Medianwert für den Touchscreen liegt. Das heißt, dass die Teilnehmer während der Benutzung des Hebels weniger frustriert waren, als bei der Bedienung des Touchscreens. Außerdem waren die Bewertungen zum Hebel innerhalb der Stichprobe konsistenter als die Bewertungen zum Touch-Interface. Inferenzstatistisch wurde allerdings auch hier kein signifikanter Unterschied zwischen dem Hebel ( $MD = 8.5$ ) und dem Touchscreen ( $MD = 24.5$ ) nachgewiesen;  $z = 1.604$ ,  $p = 0.109$ .

Tendenziell waren die Lokführer nach der Hebelbedienung mit ihrer Leistung in der Versuchsaufgabe zufriedener. Ein Indiz dafür ist die linkslastige Verteilung der Bewertungen auf der Subskala *Ausführung der Aufgabe*. Beim Touchscreen ist die Verteilung leicht rechtslastig, mit einem deutlich höheren Medianwert. Der Unterschied zwischen den Me-

dianen ( $MD$  (Hebel) = 8,  $MD$  (Touch) = 28) wurde jedoch nicht signifikant;  $z = 1.289$ ,  $p = 0.197$ .

Die Skala *körperliche Anforderung* wurde bei dem Hebel leicht höher bewertet als beim Touchscreen. Auffällig ist, dass sich die Mediane nur geringfügig unterscheiden und die Verteilung der Bewertungen zum Hebel eine größere Streuung aufweist. Beim Touch-Interface sind die Bewertungen konsistenter, was anhand der relativ kleinen Streuung der Werte zu sehen ist. Die Verteilung ist dabei symmetrisch. Der Wilcoxon-Test für Paardifferenzen ergab allerdings keinen signifikanten Unterschied bezüglich der Skala *körperliche Anforderung* zwischen dem Hebel ( $MD = 15.5$ ) und dem Touchscreen ( $MD = 9.5$ );  $z = -1.342$ ,  $p = 0.180$ .

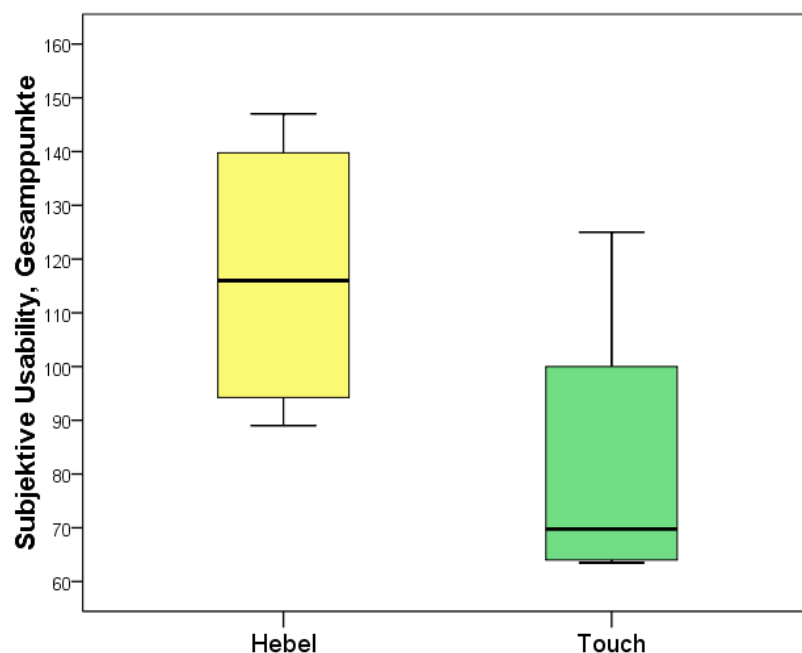


**Abbildung 9: Boxplot-Diagramm für die Verteilung der Beanspruchungshöhen der sechs Subskalen für die Bedienung von Hebel und Touchscreen**

Die Bewertungen auf der Subskala *Anstrengung* bilden linkslastige Verteilungen. Dabei zeigt sich, dass die Mediane annähernd gleich sind und die Werte eine starke Streuung aufweisen, insbesondere beim Touch-Interface. Beim Wilcoxon-Test für Paardifferenzen wurde die Hypothese, dass sich die Mediane ( $MD$  (Hebel) = 9,  $MD$  (Touch) = 10.5) auf dem Signifikanzniveau 0.05 nicht unterscheiden, vorläufig beibehalten:  $z = 1.069$ ,  $p = 0.285$ .

## 6.5 Subjektive Usability (ISONORM 9241/110-S)

Wie die *Abbildung 10* verdeutlicht, wurde der Hebel von den Lokführern insgesamt nutzerfreundlicher bewertet als das Touch-Interface: die Verteilungen überlappen sich wenig. Auch sind die unteren und oberen Quartile sowie die Mediane für die Bewertung des Hebels deutlich größer als für den Touchscreen. Die Verteilungsformen unterscheiden sich ebenso: die Verteilung der Usability-Bewertungen des Hebels ist symmetrisch und die Werte sind normalverteilt; die Verteilung des Touch-Interfaces hingegen ist eher linkslastig und der Median liegt im unteren Bereich. Der Wilcoxon-Test für Paardifferenzen ergab in der Gesamtbewertung der Usability keinen signifikanten Unterschied zwischen dem Hebel ( $MD = 116$ ) und dem Touchscreen ( $MD = 69.75$ ) auf dem Signifikanzniveau  $\alpha = 0.05$ ;  $z = -1.826$ ,  $p = 0.068$ .



**Abbildung 10: Boxplot-Diagramm für die Verteilung des Gesamtwertes der subjektiven Usability von Hebel und Touchscreen**

Auch sind die Punkte über alle sieben Usability-Faktoren beim Hebel tendenziell höher als beim Touch-Interface (*Abb. 11*). Am auffälligsten ist der Unterschied zwischen den Interfaces in der Bewertung der *Aufgabenangemessenheit*. Die Verteilungen überlappen sich wenig. Darüber hinaus sind die Bewertungen zum Hebel leicht rechtslastig verteilt (d.h. mit einer Tendenz zu höheren Werten). Die Bewertungen zum Touch-Interface weisen im Gegensatz dazu eine leicht linkslastige Verteilung auf. Allgemein sind die Bewertungen

der *Aufgabenangemessenheit* deutlich höher für den Hebel, was bedeutet, dass den Versuchspersonen der Hebel als geeigneter für die Quittierungsaufgabe erscheint, als das Touch-Interface. Auch wenn der Wilcoxon-Test keine signifikanten Unterschiede zwischen den Interfaces ( $MD$  (Hebel) = 17.5,  $MD$  (Touch) = 6.5) entdeckt hat ( $z = -1.826$ ,  $p = 0.68$ ), sollten die Ergebnisse als vorläufig anerkannt werden.

Ein weiterer Unterschied zwischen den Interfaces ist für die Dimension *Lernförderlichkeit* zu erkennen. Die Streuung ist beim Touch-Interface offensichtlich größer als beim Hebel, was bedeuten kann, dass die Bewertungen der Lernförderlichkeit für den Hebel konsistenter sind. Auch die Verteilungen unterscheiden sich: die Verteilung der Hebel-Bewertungen sieht eher symmetrisch aus, während bei dem Touchscreen eine linkslastige Verteilung vorliegt, was der Tendenz zu kleineren Werten entspricht. Die Lernförderlichkeit des Hebels wird demnach subjektiv besser bewertet als für den Touchscreen, wobei auch hier die Median-Differenz ( $MD$  (Hebel) = 19.5,  $MD$  (Touch) = 13.5) nach dem Wilcoxon-Test nicht signifikant geworden ist:  $z = -1.512$ ,  $p = 0.131$ .

Auch die *Selbstbeschreibungsfähigkeit* wurde beim Hebel besser bewertet. Im Diagramm ist eine sehr große Streuung der Werte des Touch-Interfaces zu erkennen. Auffällig ist, dass sowohl die Bewertungen des Hebels als auch die des Touchscreens eine linkslastige Verteilung aufweisen. Demnach empfanden die Versuchspersonen die Bedienung des Touchscreens als weniger intuitiv als die des Hebels. Der Wilcoxon-Test war in diesem Fall ( $MD$  (Hebel) = 13,  $MD$  (Touch) = 5.5) allerdings nicht signifikant:  $z = -1.604$ ,  $p = 0.109$ .

Auch die *Steuerbarkeit* wurde für den Hebel besser bewertet als für das Touch-Interface. Die Verteilung der Werte beim Hebel ist leicht rechtslastig und beim Touch-Interface stark linkslastig, was tendenziell für eine schlechtere Bewertung des Touchscreens spricht. Die Mediane beider Verteilungen liegen weit auseinander. Inferenzstatistisch wurde kein signifikanter Unterschied in der Steuerbarkeit zwischen dem Hebel ( $MD = 18.5$ ) und dem Touchscreen ( $MD = 8$ ) entdeckt,  $z = -1.604$ ,  $p = 0.109$ .

Das Boxplot-Diagramm veranschaulicht auch, dass der Hebel eher den Erwartungen der Lokführer entspricht: die Verteilung der Bewertungen der *Erwartungskonformität* des Hebels liegt deutlicher im oberen Bereich als die Verteilung des Touch-Interfaces. Die Verteilung beim Hebel ist fast symmetrisch, beim Touch rechtslastig, wobei der maximale Wert unter der Median-Bewertung des Hebels liegt. Die Unterschiede zwischen Hebel

( $MD = 16$ ) und Touch-Interface ( $MD = 14$ ) wurden jedoch nicht signifikant:  $z = -0.736$ ,  $p = 0.461$ .

Allgemein decken die Bewertungen der Interfaces bezüglich der Dimension *Fehlertoleranz* einen annähernd gleichen Wertebereich ab. Der Hebel wurde aber auch hier leicht besser bewertet. Die Werte weisen für beide Interfaces eine starke Streuung auf, wobei sich jedoch die Form der Verteilungen unterscheidet: die Bewertungen des Hebels sind eher normalverteilt, die Verteilung der Werte des Touchscreens hingegen ist linkslastig. Die Mediane liegen relativ weit auseinander. Aber auch hier wurde der Unterschied zwischen den Interfaces ( $MD$  (Hebel) = 16.5,  $MD$  (Touch) = 12.75) nicht signifikant:  $z = -1.000$ ,  $p = 0.317$ .

Die Subskala *Individualisierbarkeit* muss für die Auswertung ausgeschlossen werden, weil sie in diesem Kontext nicht sinnvoll ist, da die Interfaces keine Möglichkeit der Individualisierbarkeit bieten.

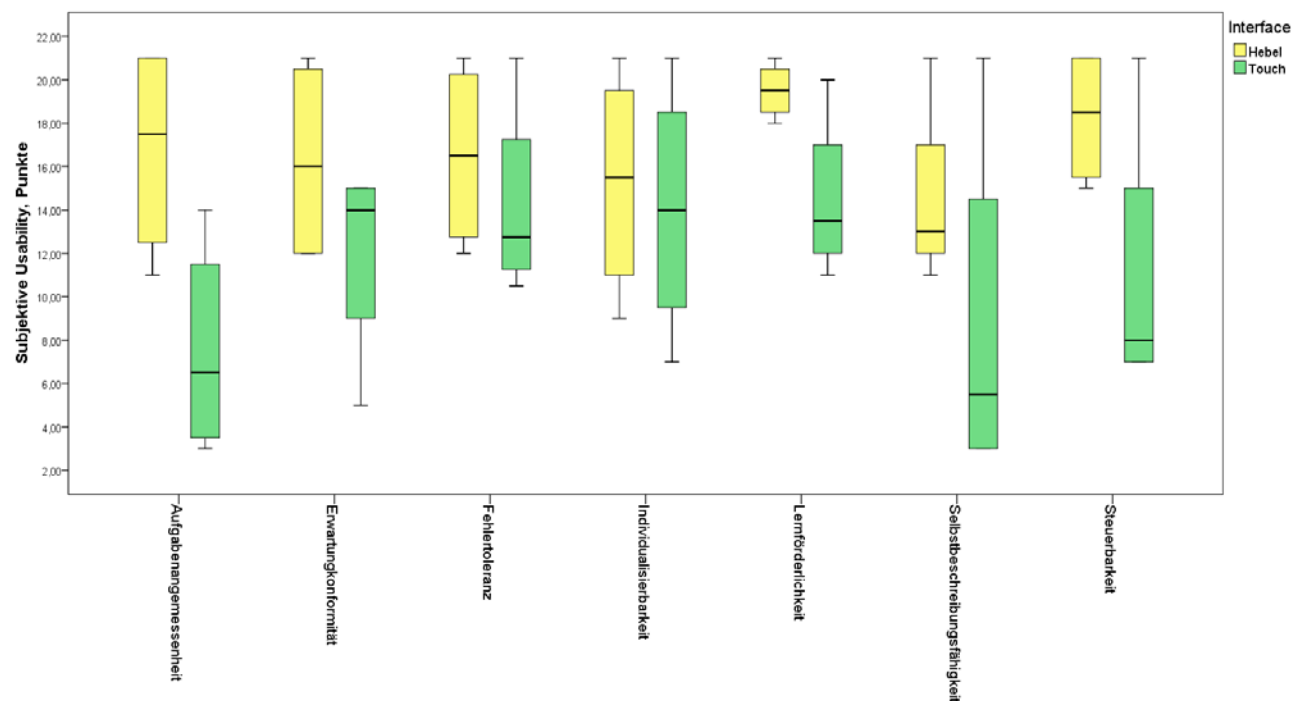


Abbildung 11: Boxplot-Diagramm für die Verteilung der Usability-Bewertungen (für die sieben Subskalen) für Hebel und Touchscreen



## 7 Diskussion

Insgesamt wurde der Hebel als weniger beanspruchend und nutzerfreundlicher von den Probanden wahrgenommen. Die Mediane auf den Subskalen *zeitliche Anforderung* und *geistige Anforderung* des NASA-TLX spiegeln das Empfinden der Probanden wider, dass die Bedienung des Touch-Interfaces im Gegensatz zum Hebel aufwändiger ist und mehr Zeit zum Nachdenken erfordert. Die höhere geistige Anforderung ist mit der Notwendigkeit verbunden, zusätzlich die Information über die Symbole aus dem Gedächtnis abzurufen. Obwohl dies bei den Probanden zu einer geringen Akzeptanz führte, ist es im Sinne der Aufgabenstellung, eine tiefere kognitive Verarbeitung zu erreichen, als positiv zu bewerten.

Bei der Quittierung mittels Touch-Interface führt neben der geistigen Anforderung auch die Handlungsausführung zu einem höheren Zeitaufwand. So erfordert sie viel visuelle Aufmerksamkeit, da über die gesamte Dauer des Zeichnens die Position des Fingers auf dem Raster überwacht werden muss. Die Handlungsausführung wird somit sehr aufwändig und bietet ein Potential zur Überforderung, ohne im direkten Zusammenhang zur kognitiven Anreicherung zu stehen.

Ein direkter Vergleich zwischen den beiden Interfaces bezüglich der erlebten Beanspruchung erscheint zudem aufgrund der unterschiedlichen kognitiven Anforderungen ohnehin schwierig. Im Vergleich zum Hebel-Interface erfordert das Touch-Interface zusätzlich eine kognitive Umkodierung der gesehenen Signale in ein entsprechendes Symbol, welches dann unter starker visueller Kontrolle eingezeichnet werden muss. Bei dem Hebel-Interface kann das wahrgenommene Signal direkt in eine entsprechende Handlung beziehungsweise Hebelstellung umgesetzt werden. Somit scheint eher ein Vergleich zwischen den jeweiligen Interface-Konzepten und dem bestehenden Interface der PZB empfehlenswert.

Um einen tatsächlichen Sicherheitsgewinn nachzuweisen, bedarf es weiterer Untersuchungen. Ob der zeitliche Aufwand durch die Quittierung den Fahrbetrieb beeinträchtigt, kann nur unter einer vergleichbaren Belastung (z. B. in einer voll funktionsfähigen Fahrsimulation) untersucht werden. Das verwendete Videomaterial, das die Zugfahrt aus Sicht des Lokführers zeigte, enthielt dahingehend kleinere Mängel. Ein Proband merkte an, dass die Geschwindigkeit sehr niedrig sei und sich bei hohen Geschwindigkeiten im Zug ein größerer Zeitdruck ergibt.

Die zwei Konzepte sind in dem Sinne sehr unterschiedlich, dass der Hebel für die Lokführer gewohnt ist, während das Touch-Interface im Führerstand eine Neuerung darstellt. Dies könnte bei den befragten Probanden eine Voreingenommenheit verursachen, die über eine längere Nutzungsdauer möglicherweise unbedeutend wird. Alle vier Probanden gaben jedoch an, mit der Bedienung von Touchscreens vertraut zu sein, sodass Schwierigkeiten in der Benutzung nicht auf mangelnde Erfahrung zurückzuführen sein sollte. Trotzdem ist noch zu untersuchen, ob sich eine längere Eingewöhnungsphase auf die Bewertung des Touch-Interfaces auswirkt. Hinzukommt, dass sich die generelle Einstellung bezüglich Technik hier widerspiegeln könnte. Es wäre daher ratsam, in Folgeuntersuchungen derlei Einflüsse zu berücksichtigen und ggf. mittels Fragebogen vor der Untersuchung zu erfassen. Ebenso ist es möglich, dass sich die Länge der Berufserfahrung auf die Akzeptanz und den Trainingsaufwand für ein neues System auswirkt.

Im Interview wurden die positiven und negativen Aspekte für beide Interfaces gemeinsam erhoben. Im diesem Zusammenhang ist nicht klar, ob das Touchpad nicht eine positive Anmerkung bekommen hat, weil die Fragestellung den Probanden nicht ausreichend aufgefordert hat, darüber nachzudenken.

Ein Kritikpunkt zur Auswertung der qualitativen Daten betrifft die mangelnde Interrater-Reliabilität. Die Kategorienbildung der qualitativen Rohdaten erfolgt üblicherweise durch mehrere Personen. Darauf wurde jedoch im Rahmen dieser Projektarbeit aufgrund begrenzter Ressourcen verzichtet. Allerdings erzielten die, durch zwei unabhängige Bewerter angefertigten Kategorien für das Laute Denken und das Interview große inhaltliche Übereinstimmungen.

Insbesondere für die Interpretation statistischer Daten ist die erhobene Stichprobe zu gering. Für die Bewertung von subjektiver Beanspruchung und subjektiver Usability können demnach nur Trends abgebildet werden, die jedoch im Einklang mit den qualitativen Ergebnissen stehen. So spiegelt sich beispielsweise die Aussage, das Touch-Interface sei zu aufmerksamkeitsbindend und zeitaufwändig, in einer höheren Bewertung des Touch-Interfaces auf den Subskalen *zeitliche Anforderung* und *geistige Anforderung* wider.

Die Ergebnisse der Evaluation können lediglich als richtungsweisende Exploration auf dem Weg zur Entwicklung eines gut ausgereiften Interfaces gesehen werden. Durch den frühen Einsatz von Prototypen konnten jedoch kostengünstig Erkenntnisse für weitere Iterationen gesammelt werden, bevor aufwendigere Untersuchung durchgeführt werden, die

Aussagen über den kritischen Aspekt der Sicherheit erlauben. Für diese erscheint eine umfassendere Einbindung der Zielgruppe in Form von Befragungen oder einer intensiven Beschäftigung mit detaillierten Handlungsabläufen eines Triebfahrzeugführers, insbesondere in sehr frühen Phasen der Konzeptentwicklung, notwendig.

## 8 Ausblick

Die Untersuchungen der prototypischen Interfaces ergaben viele interessante Einblicke und eröffneten die Möglichkeit, weitere Untersuchungen anzuschließen. Im Folgenden soll ein kleiner Ausblick für Verbesserungsvorschläge der Interfaces, Empfehlungen weiterer Untersuchungsmethoden und potentielle Forschungsfragen erfolgen.

Das Touchpad gab am meisten Anlass zur Kritik und beinhaltet sehr viel Verbesserungspotential. Ein häufig genannter Kritikpunkt betrifft dabei die Gestaltung des Eingabefensters. Die Quittierungshandlung innerhalb eines vorgegebenen Punkterasters erforderte sehr viel visuelle Aufmerksamkeit und verhinderte, dass die Lokführer die Strecke durchgehend beobachten konnten. Desweiteren benötigte die Ausführung sehr viel Zeit. In weiteren Entwicklungsschritten sollte man deshalb die Gestaltung des interaktiven Elementes so gestalten, dass eine Eingabe ohne die längere Blickzuwendung möglich ist.

Um den zeitlichen Aufwand zu minimieren, könnte die Symbolik vereinfacht werden. Zum Beispiel wäre es denkbar, statt eines Buchstaben, einen einfachen Strich zu verwenden.

Für die prototypische Umsetzung der Interfaces wurde bisher kein ausgereiftes Feedbacksystem entwickelt. Während der Testung wurde dies von vielen Versuchspersonen als wichtiger Verbesserungsvorschlag bemerkt. Insbesondere ein haptisches Feedback, etwa in Form eines Vibrationsimpulses, sollte bei der Quittierungshandlung als Bestätigung einer erfolgreichen Eingabe verwendet werden. Denkbar wäre auch ein optisches Feedback, wie beispielsweise eine grüne Einfärbung der Eingabesymbole, wenn diese von der Software erkannt wurden oder ein akustisches Signal, welches eine abgeschlossene Eingabehandlung bestätigt.

Auch das Hebelsystem bietet viele Möglichkeiten zur weiteren Erforschung. Bisher enthält das Hebel-Interface drei Informationsquellen, anhand derer die jeweilige Zuordnung der Vorsignale erleichtert werden soll: unterschiedliche LEDs, die verschiedenen Hebelstellungen und die Beschriftung der jeweiligen Hebeleinstellungen. Hier könnte in weiteren

Forschungsschritten untersucht werden, inwiefern sich das Weglassen einer oder mehrerer Hilfestellungen auf die Bedienbarkeit des Interfaces auswirkt. Es ist denkbar, dass einzig eine Veränderung der Hebelstellung als Quittierung der jeweiligen Vorsignale ausreicht und eher zur kognitiven Anreicherung dieser Handlung beitragen könnte als das bisherige Konzept.

Für die Durchführung weiterer Forschungsarbeiten empfiehlt sich die Verwendung einer realitätsgetreuen Untersuchungsumgebung, damit die Benutzung der Interfaces innerhalb des angestrebten Arbeitsumfeldes untersucht werden kann. So könnte beispielsweise der Einsatz von Simulatoren dazu dienen, den Realitätsgrad der Untersuchung zu erhöhen. Zusätzlich ermöglicht ein Zugsimulator, die Komplexität der Arbeitsaufgabe genauer abzubilden sowie die Quittierungshandlung in den gesamten Handlungskontext einzubetten. Dabei können arbeitsbedingte Aspekte wie Monotonie und Ermüdung stärker berücksichtigt werden.

Außerdem könnte für eine weitergehende Untersuchung der kognitiven Verarbeitung der Einsatz verschiedener Methoden berücksichtigt werden. So könnten mit Hilfe von Blickbewegungserfassungen die Gestaltung der Interfaces evaluiert werden. Parameter wie beispielsweise die Anzahl an Fixationen und die Verweildauer (*dwell time*) könnten Hinweise auf das Ausmaß der Bindung von visueller Aufmerksamkeit geben.

Durch die zusätzliche Erhebung physiologischer Daten könnten weitere Ergebnisse zur mentalen Beanspruchung gewonnen werden. Das ist insofern wichtig, weil das subjektive Empfinden von Beanspruchung nicht den physiologischen Reaktionen entsprechen muss.

Weitere Messungen wie z.B. die Erfassung des Situationsbewusstseins sind von Interesse, da sie das Verständnis des Lokführers bezüglich der aktuellen Situation überprüfen. So lässt sich ableiten, ob das Interface zu komplex ist, um die jeweilige Situation adäquat verstehen zu können. Eine mögliche Anwendung stellt die *Situation Awareness Global Assessment Technique* (SAGAT) dar.

Außerdem ist es empfehlenswert ein Experiment mit einer Kontrollgruppe bzw. die Erhebung der Daten der bestehenden Quittierungshandlung als Baseline zu realisieren. Dazu könnten die Vergleiche unter verschiedenen kontrollierten Bedingungen sowie für unterschiedliche Geschwindigkeiten, Tages- und Jahreszeiten, verschiedene Wetterbedingungen und unterschiedliche Vorsignal- und Loktypen durchgeführt werden.

Für einen besseren Umgang mit den erhobenen Daten und einen statistisch relevanten Informationsgewinn sollte auch die Stichprobe größer sein. Die Fragestellungen bezüglich folgender Gebiete könnten im Sichtfeld der weiteren Untersuchungen stehen: die Akzeptanz von Lokführern, Interferenzen mit anderen Fahraufgaben, Trainingseffekte und die Integration eines neuen Interfaces im Cockpit.

## A. Anhang – Demografischer Fragebogen

Code: \_\_\_\_\_

(ersten beiden Anfangsbuchstaben Mutter + ersten beiden Anfangsbuchstaben Vater + ersten beiden Ziffern Geburtsdatum, z.B. SUMA03)

**Zunächst würden wir Sie bitten ein paar kurze Fragen zu Ihrer Person und Ihrem Beruf zu beantworten. Alle Angaben sind anonymisiert.**

### Angaben zur eigenen Person

Alter: \_\_\_\_\_

Geschlecht:

- ☐ männlich
- ☐ weiblich

Händigkeit:

- ☐ linkshändig
- ☐ rechtshändig
- ☐ beidhändig

Sehschwäche:

- ☐ ja
- ☐ nein

Haben Sie bereits Erfahrungen im Umgang mit Touchscreens (Smartphones, Tablets, etc.)?

- ☐ ja
- ☐ nein

### Angaben zum Beruf

Wie lange üben Sie diesen Beruf bereits aus? \_\_\_\_\_

Einsatzgebiet (wo innerhalb von Deutschland werden Sie größtenteils eingesetzt?):

\_\_\_\_\_

Fahrzeug:

- ☐ S- Bahn
- ☐ Regionalbahn
- ☐ ICE
- ☐ sonstige: \_\_\_\_\_

Durchschnittliche Fahrdauer (pro Schicht, in Stunden): \_\_\_\_\_

Anzahl der Pausen (pro Schicht): \_\_\_\_\_

Länge der Pausen (in Minuten): \_\_\_\_\_

Arbeitszeiten (Schichten, Mehrfachauswahl möglich):

- ☐ vorwiegend morgens
- ☐ vorwiegend tagsüber
- ☐ vorwiegend nachts
- ☐ ausgeglichen

Wann war Ihre letzte Schulung/ Weiterbildung?

\_\_\_\_\_

Wie oft erhalten sie Schulungen/ Weiterbildungen?

\_\_\_\_\_

Fahren Sie auch außerhalb von Deutschland?

- ☐ ja
- ☐ nein

Wenn ja, wo?

\_\_\_\_\_

## B. Anhang – Einführungs und Begrüßungstext

*Technische Universität Berlin*



Institut für Psychologie und  
Arbeitswissenschaft



Fachgebiet

Mensch-Maschine-Systeme

Prof. Dr.-Ing. Matthias Rötting

Willkommen zum Experiment:

### **Beurteilung von Schnittstellen zur Quittierung des Vorsignals auf Basis der PZB**

#### *Einführung*

Vielen Dank, dass Sie an unserer Untersuchung teilnehmen und uns dadurch bei unserem Projekt helfen. Die Teilnahme an der Untersuchung ist selbstverständlich freiwillig. Die erhobenen Daten werden vertraulich behandelt, ohne Bezug zu Ihrer Person ausgewertet, ausschließlich für statistische Zwecke genutzt und keinesfalls an Dritte weitergegeben.

#### *Zweck der Untersuchung*

Die Technische Universität Berlin möchte im Auftrag des Deutschen Instituts für Luft- und Raumfahrt (DLR) ein neues Bedienelement für die Quittierung der Vorsignale auf Basis der PZB entwickeln. Dazu werden wir Ihnen gleich zwei verschiedene Varianten präsentieren. Uns interessiert, welches Sie in der Nutzung bevorzugen. Zu diesem Zweck möchten wir Ihre Meinung wissen. Wichtig ist nur Ihre eigene Meinung, deswegen gibt es keine richtigen oder falschen Antworten. Es wird hierbei ausschließlich die Gebrauchstauglichkeit der Systeme getestet. Wir bitten Sie, darauf zu achten, so genau und ehrlich wie möglich zu antworten.



## *Überblick*

Im Folgenden werden Sie Videoausschnitte einer Bahnfahrt sehen, bei denen Vorsignale zu quittieren sind. Dazu stehen Ihnen zwei Bedienelemente zur Verfügung, bei denen die Quittierung unterschiedlich realisiert werden soll. Zunächst erlernen Sie die zwei unterschiedlichen Eingaben zur Bestätigung der Vorsignale. Anschließend erfolgt ein kurzer Probedurchlauf. Schließlich werden Sie eine digitale Teststrecke befahren und dabei auftretende Vorsignale mit dem neuen Bedienelement bestätigen. Während der Durchführung möchten wir die Technik des Lauten Denkens anwenden. Dabei geht es darum, Ihre Gedanken zu verbalisieren, sodass wir nachempfinden können, wie Sie die Anwendung erleben. Dazu bitten wir Sie, frei heraus zu sagen, was Sie denken. Wichtig ist, dass Sie alle Gedanken nennen, auch jene, die Ihnen irrelevant erscheinen. Ihre Meinung zu den Bedienelementen sowie Ihre Beanspruchung werden wir zusätzlich mittels Fragebogen sowie zum Abschluss mittels Interview erfassen.

Um eine folgende Analyse zu erleichtern und mögliche Datenfehler zu umgehen, möchten wir ihre Erzählungen aufzeichnen.

Der gesamte Ablauf wird etwa 2 Stunden betragen.

Sie erhalten für Ihre Teilnahme 20 €.

Sie können Ihre Teilnahme an dem Versuch jederzeit abbrechen, ohne Angabe von Gründen und ohne dass Ihnen daraus Nachteile entstehen.

## C. Anhang – Lautes Denken: Ablauf

### Ablauf Lautes Denken

1. Vorbereitung des Experiments:
  - Laborsetting vorbereiten
  - Technik einstellen (Kamera, Tonaufzeichnung)
  - Probedurchlauf, um Technik zu checken
  - Instruktionen bereitlegen
2. Begrüßung:
  - Smalltalk
  - angenehme Atmosphäre schaffen
  - Sicherheit vermitteln
3. Einführung:
  - Ziel und Dauer der Untersuchung erläutern
  - Verfahren und ggf. Technik erklären
  - eventuell kurzes Video aus vorherigen Untersuchungen vorführen (damit die Methode des Lauten Denkens verständlich wird)
  - Einverständniserklärung unterschreiben lassen
4. Probedurchlauf:
  - Bearbeitung einer Übungsaufgabe, um Versuchsperson an das laute Denken zu gewöhnen
  - Feedback geben (z.B. „Das haben Sie schon sehr gut gemacht. Trauen Sie sich ruhig noch mehr Gedanken zu äußern.“)
5. Aufgabe:
  - Instruktion aushändigen
  - Aufgabe sollte möglichst nicht vom Versuchsleiter erklärt werden
6. Durchführung der Aufgabe:
  - Versuchsperson bearbeitet Aufgabe und spricht dabei laut auftretende Gedanken aus
  - Versuchsleiter kann die Testperson zum Sprechen ermutigen, wenn diese lange nichts gesagt hat
  - Beobachter notiert Auffälligkeiten
7. Abschluss:
  - ggf. Interview, in dem bestimmte Äußerungen genauer erläutert werden sollen

- Fragebögen aushändigen
- abschließendes Interview
- Geld aushändigen & in Liste unterschreiben lassen (Geld erhalten)
- Verabschiedung

#### 8. Auswertung

- Transkription der Notizen, Video- und Soundaufnahmen
- Kategorisieren der aufgetretenen Probleme, Bewertung der jeweiligen Schwere des Problems, Kreuzprodukt aus Anzahl der Nennungen und Schweregrad bilden und daraus eine Rangliste erstellen
- Analysieren weiterer Notizen

## D. Anhang – Lautes Denken: Text zum Vorlesen

Ich werde jetzt ein Interface benutzen und den Versuchsleiter daran teilhaben lassen, wie es auf mich wirkt. Ich werde versuchen, während der Nutzung laut zu denken.

Ich werde laut reden und meinen Versuchsleitern mitteilen, welche Ziele ich mir setze, was mir an dem Interface gut und was mir nicht so gut gefällt, was mich wundert oder irritiert.

Dieses laute Denken wird sich vielleicht etwas ungewohnt anfühlen, aber das ist vollkommen okay. Ich werde einfach in dem Tempo arbeiten und reden, in dem ich mich am Wohlsten fühle.

Ich kann ganz entspannt sein, denn nicht ich werde hier bewertet, sondern das Interface. Es gibt also kein richtig oder falsch, sondern nur mein persönliches Empfinden.

Mein Feedback wird dabei helfen, das Interface weiter zu verbessern.

## E. Anhang – Instruktionen für die Versuchsperson

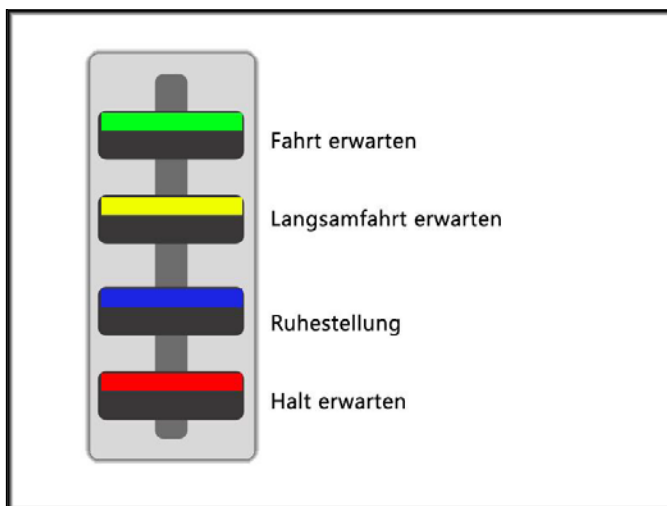
### Aufgabe

Während des Versuchs ist es Ihre Aufgabe, die von Ihnen erkannten **Vorsignale** eines Videoausschnitts mithilfe des Interfaces zu quittieren. Sprechen Sie dabei bitte stets laut Ihre Gedanken, Absichten und Vorstellungen aus.

#### Beschreibung des Interfaces

Es handelt sich um einen **Hebel**, welcher in vier unterschiedliche Stellungen gebracht werden kann. Ist ein Hebel in einer der beschriebenen Positionen, leuchtet eine entsprechende Lampe auf.

Bitte beachten Sie dabei folgende Zuordnung:



Hebelpositionen:

- |   |                                    |
|---|------------------------------------|
| • Fahrt erwarten                        | → Hebel ganz nach vorne schieben   |
| • Geschwindigkeitsbeschränkung erwarten | → Hebel leicht nach vorne schieben |
| • Ruhestellung                          | → Hebel in der Mittelposition      |
| • Halt erwarten                         | → Hebel nach hinten ziehen         |

Nachdem Sie ein Vorsignal „passiert“ haben, geben Sie bitte die entsprechende Hebelstellung ein. Dafür haben Sie **4 Sekunden** Zeit. Den Hebel bringen Sie bitte erst wieder in Ruhestellung (Mittelposition, blaue Lampe), bevor Sie auf das nächste Vorsignal reagieren möchten!

Also: Hebeleinstellung für Vorsignal 1 → Sehen des zweiten Vorsignals → Ruhestellung → Hebeleinstellung für Vorsignal 2.

## Aufgabe

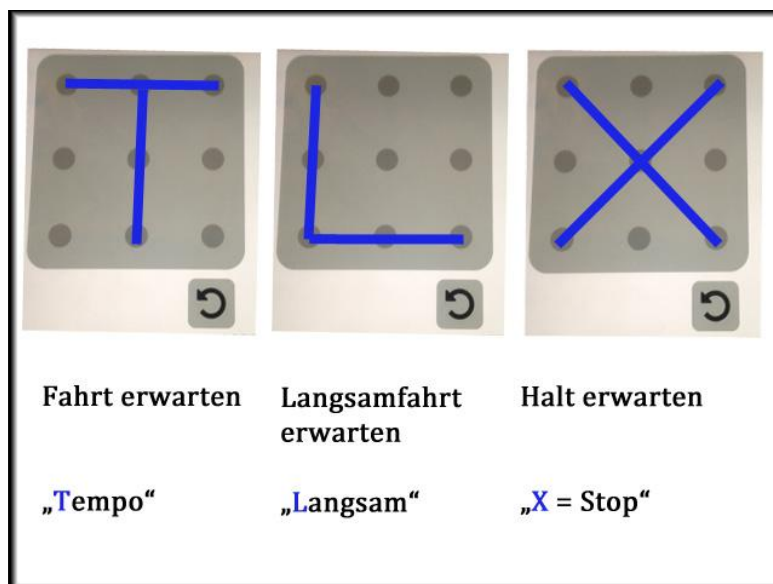
Während des Versuchs ist es Ihre Aufgabe die von Ihnen erkannten **Vorsignale** eines Videoausschnitts mithilfe der Interfaces zu quittieren. Sprechen Sie dabei bitte stets laut Ihre Gedanken, Absichten und Vorstellungen aus.

### Beschreibung des Interfaces

Bei diesem Interface handelt es sich um ein Touchdisplay mit einem **3 x 3 Punkteraster**. Auf diesem können Sie mit dem Finger die entsprechenden Vorsignale einzeichnen. Dabei sollten Sie sich an dem Raster orientieren und nicht über den Rand zeichnen.

Zur Quittierung der verschiedenen Vorsignale sollen drei verschiedene Buchstaben (**L**, **X** und **T**) gezeichnet werden.

Bitte beachten Sie dabei folgende Zuordnungen:



Die Beschreibungen „Tempo“, „Langsam“ und „X = Stop“ sollen Ihnen lediglich bei der Zuordnung der Buchstaben zu den jeweiligen Vorsignalen helfen.

Sollten Sie sich dabei vertun, können Sie den „Zurück“-Button anklicken. Dann verschwindet Ihre Eingabe und Sie können erneut beginnen.

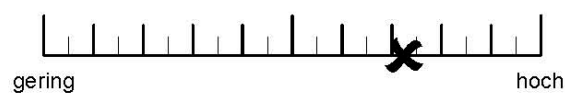
Ihr gezeichnetes Symbol wird nach **4 Sekunden** vom Bildschirm verschwinden.

## F. Anhang – NASA-TLX

### Beanspruchungshöhe

Geben Sie jetzt bitte an, wie hoch die Beanspruchung in den einzelnen Dimensionen war. Markieren sie dazu auf den folgenden Skalen bitte, in welchem Maße Sie sich in den sechs genannten Dimensionen von der Aufgabe beansprucht oder gefordert gesehen haben:

Beispiel:



### Geistige Anforderungen

Wie viel geistige Anstrengung war bei der Informationsaufnahme und bei der Informationsverarbeitung erforderlich (z.B. Denken, Entscheiden, Rechnen, Erinnern, Hinsehen, Suchen ...)? War die Aufgabe leicht oder anspruchsvoll, einfach oder komplex, erfordert sie hohe Genauigkeit oder ist sie fehlertolerant?



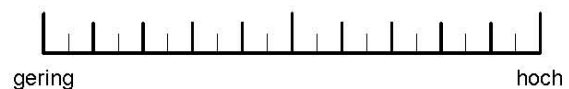
### Körperliche Anforderungen

Wie viel körperliche Aktivität war erforderlich (z.B. ziehen, drücken, drehen, steuern, aktivieren ...)? War die Aufgabe leicht oder schwer, einfach oder anstrengend, erholsam oder mühselig?



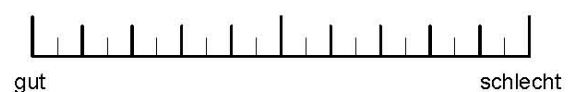
### Zeitliche Anforderungen

Wie viel Zeitdruck empfanden Sie hinsichtlich der Häufigkeit oder dem Takt mit dem Aufgaben oder Aufgabenelemente auftraten? War die Abfolge langsam und geruhsam oder schnell und hektisch?



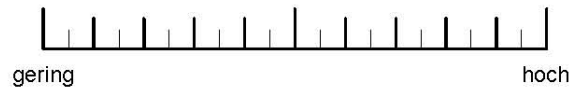
### Ausführung der Aufgaben

Wie erfolgreich haben Sie ihrer Meinung nach die vom Versuchsleiter (oder Ihnen selbst) gesetzten Ziele erreicht? Wie zufrieden waren Sie mit Ihrer Leistung bei der Verfolgung dieser Ziele?



## Anstrengung

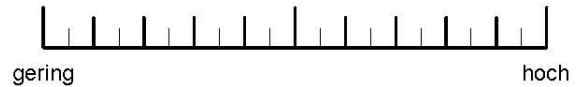
Wie hart mussten Sie arbeiten, um Ihren Grad an Aufgabenerfüllung zu erreichen?



gering hoch

## Frustration

Wie unsicher, entmutigt, irritiert, gestresst und verärgert (versus sicher, bestätigt, zufrieden, entspannt und zufrieden mit sich selbst) fühlten Sie sich während der Aufgabe?



gering hoch

Kontrollieren Sie bitte, ob Sie zu allen Fragen Angaben gemacht haben. Bei Unklarheiten wenden Sie sich bitte an die anwesenden Versuchsleiter.



## G. Anhang – ISONORM 9241/110-S

### Ein Hinweis zur Beantwortung des Beurteilungsbogens:

Im folgenden Fragebogen werden die Anforderungen der Norm über Beschreibungen konkretisiert. Diese Beschreibungen weisen immer folgende Form auf:

### Beispiel:

...Das Interface	---	--	-	- / +	+	++	+++	... Das Interface
ist schlecht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	ist gut.

Im ersten Beispiel wird danach gefragt, wie gut bzw. wie schlecht das Interface ist. Die Benutzerin oder der Benutzer beurteilt in diesem Fall das Interface zwar als gut, sieht jedoch noch Verbesserungsmöglichkeiten.

Am besten bearbeiten Sie den Beurteilungsbogen, während Sie das zu bewertende System vor sich haben. Dadurch haben Sie die Möglichkeit, bei der Beantwortung der einzelnen Fragen die eine oder andere Sache noch einmal zu überprüfen.

Füllen Sie bitte den Beurteilungsbogen äußerst sorgfältig aus und lassen Sie keine der Fragen aus!

	Das Interface ...	--- -- - -/+ + ++ +++	Das Interface ...
aa1	bietet nicht alle Funktionen, um die anfallenden Aufgaben effizient zu bewältigen.	○ ○ ○ ○ ○ ○ ○	bietet alle Funktionen, um die anfallenden Aufgaben effizient zu bewältigen.
aa2	erfordert überflüssige Eingaben.	○ ○ ○ ○ ○ ○ ○	erfordert keine überflüssigen Eingaben.
aa3	ist schlecht auf die Anforderungen der Arbeit zugeschnitten.	○ ○ ○ ○ ○ ○ ○	ist gut auf die Anforderungen der Arbeit zugeschnitten.
sb1	liefert in unzureichendem Maße Informationen darüber, welche Eingaben zulässig oder nötig sind.	○ ○ ○ ○ ○ ○ ○	liefert in zureichendem Maße Informationen darüber, welche Eingaben zulässig oder nötig sind.
sb2	bietet auf Verlangen keine situationsspezifischen Erklärungen, die konkret weiterhelfen.	○ ○ ○ ○ ○ ○ ○	bietet auf Verlangen situationsspezifische Erklärungen, die konkret weiterhelfen.
sb3	bietet von sich aus keine situationsspezifischen Erklärungen, die konkret weiterhelfen.	○ ○ ○ ○ ○ ○ ○	bietet von sich aus situationsspezifische Erklärungen, die konkret weiterhelfen.
ek1	erschwert die Orientierung durch eine uneinheitliche Gestaltung.	○ ○ ○ ○ ○ ○ ○	erleichtert die Orientierung durch eine einheitliche Gestaltung.

	Das Interface ...	--- -- - -/+ + ++ +++	Das Interface ...
ek2	informiert in unzureichendem Maße über das, was es gerade macht.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	informiert in ausreichendem Maße über das, was es gerade macht.
ek3	lässt sich nicht durchgehend nach einem einheitlichen Prinzip bedienen.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	lässt sich durchgehend nach einem einheitlichen Prinzip bedienen.
lf1	erfordert viel Zeit zum Erlernen.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	erfordert wenig Zeit zum Erlernen.
lf2	erfordert, dass man sich viele Details merken muss.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	erfordert nicht, dass man sich viele Details merken muss.
lf3	ist schlecht ohne fremde Hilfe oder Handbuch erlernbar.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	ist gut ohne fremde Hilfe oder Handbuch erlernbar.
sk1	erzwingt eine unnötig starre Einhaltung von Bearbeitungsschritten.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	erzwingt keine unnötig starre Einhaltung von Bearbeitungsschritten.
sk2	ermöglicht keinen leichten Wechsel zwischen einzelnen Menüs oder Masken.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	ermöglicht einen leichten Wechsel zwischen einzelnen Menüs oder Masken.

	Das Interface ...	--- -- - -/+ + ++ +++	Das Interface ...
sk3	erzwingt unnötige Unterbrechungen der Arbeit.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	erzwingt keine unnötigen Unterbrechungen der Arbeit.
ft2	erfordert bei Fehlern im Großen und Ganzen einen hohen Korrekturaufwand.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	erfordert bei Fehlern im Großen und Ganzen einen geringen Korrekturaufwand.
ft3	gibt keine konkreten Hinweise zur Fehlerbehebung.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	gibt konkrete Hinweise zur Fehlerbehebung.
lk1	lässt sich schwer erweitern, wenn für mich neue Aufgaben entstehen.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	lässt sich leicht erweitern, wenn für mich neue Aufgaben entstehen.
lk2	lässt sich schlecht an meine persönliche, individuelle Art der Arbeitserledigung anpassen.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	lässt sich gut an meine persönliche, individuelle Art der Arbeitserledigung anpassen.
lk3	lässt sich - im Rahmen ihres Leistungsumfangs - von mir schlecht für unterschiedliche Aufgaben passend einrichten.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	lässt sich – im Rahmen ihres Leistungsumfangs - von mir gut für unterschiedliche Aufgaben passend einrichten.

## H. Anhang – Interview

Vielen Dank, dass Sie sich die Zeit genommen haben, um an unserem Versuch teilzunehmen. Zum Schluss würde ich Ihnen gerne noch einige Fragen zu den beiden Systemen stellen, mit denen sie gerade gearbeitet haben.

Notizen Interviewer: *(notieren, welches System welchem Prototypen entsprach + Vpn- Code)*

---

---

---

---

### Fragen:

1. Welches System hat Ihnen **besser** gefallen? Warum?

---

---

---

---

2. Welches System fanden Sie **komplizierter**?/ Welches System konnten Sie leichter bedienen?

---

---

---

---

3. Bei welchem System fiel es Ihnen leichter die entsprechenden Vorseignale zu **lernen**/ anzuwenden?

---

---

---

---

4. Gab es an irgendeiner Stelle **Probleme**? Wenn ja, welche?

---

---

---

---

5. Haben Sie eventuell irgendwelche **Verbesserungsvorschläge**?

---

---

---

---

6. Mit welchem System würden Sie sich im Berufsalltag **sicherer** fühlen? Warum?

---

---

---

---

7. Finden sie vielleicht eines der beiden Systeme **besser als die aktuelle** Quittierungshandlung? Wenn ja, warum? Was gefällt Ihnen daran besser?

---

---

---

---

8. Halten Sie die Systeme für **angemessen**, um damit die Vorsignale abzufragen?

---

---

---

---

9. Haben Sie das Gefühl, dass Sie mit den Systemen die **Vorsignale bewusster verarbeiten**?

---

---

---

---

10. Haben Sie anhand der Interfaces Analogien (**Metaphern**) zu bekannten Verkehrsabläufen (z.B. Autofahrt, Zugfahrt) gesehen? Halten sie diese für angemessen und verständlich? angemessen?

---

---

---

---

11. Haben Sie bei der Benutzung mit dem Touch Display im Gedächtnis, welche Wörter Sie für die 3 Buchstaben eingesetzt haben? Wenn ja, welche?

---

---

---

---

## Tabellenverzeichnis

SYSTEME MIT PUNKTFÖRMIGER ZUGBEEINFLUSSUNG IN ANDEREN LÄNDERN	8
---	---



## Abbildungsverzeichnis

ABBILDUNG 1: BETRIEBSPROGRAMM DER PZB .....	7
ABBILDUNG 2: HEBEL-KONZEPT .....	14
ABBILDUNG 3: HEBEL-PROTOTYP.....	15
ABBILDUNG 4: TOUCH-INTERFACE-KONZEPT .....	15
ABBILDUNG 5: TOUCH-INTERFACE-PROTOTYP .....	16
ABBILDUNG 6: VERSUCHSAUFBAU .....	19
ABBILDUNG 7: BOXPLOT-DIAGRAMM FÜR DIE VERTEILUNG DER FEHLERRATE BEI DER NUTZUNG DES HEBELS UND DES TOUCHSCREENS .....	26
ABBILDUNG 8: BOXPLOT-DIAGRAMM FÜR DIE VERTEILUNG DES GESAMTEN WERTES ZUR SUBJEKTIVEN BEANSPRUCHUNG BEI DER NUTZUNG DES HEBELS UND TOUCHSCREENS .....	27
ABBILDUNG 9: BOXPLOT-DIAGRAMM FÜR DIE VERTEILUNG DER BEANSPRUCHUNGSHÖHEN DER SECHS SUBSKALEN FÜR DIE BEDIENUNG VON HEBEL UND TOUCHSCREEN .....	29
ABBILDUNG 10: BOXPLOT-DIAGRAMM FÜR DIE VERTEILUNG DES GESAMTWERTES DER SUBJEKTIVEN USABILITY VON HEBEL UND TOUCHSCREEN.....	30
ABBILDUNG 11: BOXPLOT-DIAGRAMM FÜR DIE VERTEILUNG DER USABILITY- BEWERTUNGEN (FÜR DIE SIEBEN SUBSKALEN) FÜR HEBEL UND TOUCHSCREEN.....	32

## Literaturverzeichnis

- Allgemeines Eisenbahngesetz (AEG) - Gesetze im Internet.* (2005). Abgerufen am 12.02.2014 von [http://www.gesetze-im-internet.de/bundesrecht/aeg\\_1994/gesamt.pdf](http://www.gesetze-im-internet.de/bundesrecht/aeg_1994/gesamt.pdf).
- Anlage 4 TfV - Buzer.de.* (2011). Abgerufen am 13.02.2014 von <http://www.buzer.de/gesetz/9710/a171140.htm>.
- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. 7. Auflage. Berlin, Heidelberg: Springer.
- Buzer.de.* Abgerufen von <http://www.buzer.de/gesetz/9710/a171140.htm>.
- DB Netz AG.* (2012). Abgerufen am 12.02.2014 von [http://www.db-netz.de/fahrweg-de/start/technik/strukturseite/etcs\\_allgemeine\\_informationen.html](http://www.db-netz.de/fahrweg-de/start/technik/strukturseite/etcs_allgemeine_informationen.html).
- DIN EN ISO 9241-110. (2008). *DIN EN ISO 9241-110: Ergonomie der Mensch-System-Interaktion - Teil 110: Grundsätze der Dialoggestaltung (ISO 9241-110)*. Berlin: Beuth.
- Eisenbahnjournal Zughalt.de. (2011). *Eisenbahnjournal Zughalt.de*. Abgerufen am 13.02.2014 von <http://www.zughalt.de/2011/02/pzb-was-ist-das-eigentlich>.
- Fahrmeir, L., Künstler, R., Pigeot, I., & Tutz, G. (2007). *Statistik. Der Weg zur Datenanalyse*. 6. Auflage. Berlin, Heidelberg, New York: Springer.
- FOCUS Online.* Abgerufen am 12.02.2014 von [http://www.focus.de/finanzen/news/so-viele-fahrgaeste-wie-noch-nie-deutsche-bahn-faehrt-erneut-rekordgewinn-ein\\_aid\\_945396.html](http://www.focus.de/finanzen/news/so-viele-fahrgaeste-wie-noch-nie-deutsche-bahn-faehrt-erneut-rekordgewinn-ein_aid_945396.html).
- Gesetze im Internet.* Abgerufen von <http://www.gesetze-im-internet.de/tfv/BJNR070510011.html>.
- Gesetze im-Internet.* Abgerufen von [http://www.gesetze-im-internet.de/bundesrecht/aeg\\_1994/gesamt.pdf](http://www.gesetze-im-internet.de/bundesrecht/aeg_1994/gesamt.pdf).
- Grundlagen der Stellwerkstechnik.* (2004). Abgerufen am 14.02.2014 von [www.stellwerke.de](http://www.stellwerke.de): [http://www.stellwerke.de/grund/seite1\\_8.html](http://www.stellwerke.de/grund/seite1_8.html).

- Heuer, J. (2003). Expertenevaluation. In S. V. Heinsen, *Usability praktisch umsetzen* (S. 116-135). München.
- Mayring, P. (2010). *Qualitative Inhalstanalyse - Grundlagen und Techniken (11th ed.)*. Weinheim und Basel: Beltz.
- PZB 90. (2011). Abgerufen am 14.02.2014 von <http://www.tf-ausbildung.de/BahnInfo/pzb90.htm>.
- PZB 90 Betriebsprogramm.PNG - Wikimedia Commons. (2013). Abgerufen am 14.02.2014 von [http://commons.wikimedia.org/wiki/File:PZB\\_90\\_Betriebsprogramm.PNG](http://commons.wikimedia.org/wiki/File:PZB_90_Betriebsprogramm.PNG).
- Schmid, M., & Maier, T. (2009). Ergonomische Interfacegestaltung - Funktionalität versus Schönheit? In *Der Mensch im Mittelpunkt technischer Systeme, 8. Berliner Werkstatt Mensch-Maschine-Systeme (Volltext auf beiliegender CD)* (S. 143-144). Berlin.
- Triebfahrzeugführerscheinverordnung - TfV - Gesetze im Internet. (2011). Abgerufen am 13.02.2014 von <http://www.gesetze-im-internet.de/tfv/BJNR070510011.html>.
- Was ist ein Punktförmiges Zugbeeinflussungssystem kurz PZB. (2013). Abgerufen am 14.02.2014 von <http://ts20xxcommunity.de/index.php?page=Thread&threadID=41>.
- Wegener, M. (2009). *Informationen und Diskussionsbeiträge zum Schienenverkehr*. Abgerufen am 13.02.2014 von <http://www.marco-wegener.de/technik/indusi.htm>.